Alanya Akademik Bakış Dergisi
Yıl: 2024, C:8, S:1, s.1-14

Alanya Academic Review Journal
Year: 2024, Vol:8, No:1, p.1-14

Estimation and Comparative Analysis of Generalized Ordered Logit and Multinomial Logit Models*

(Research Article)

Genelleştirilmiş Sıralı Logit ve Multinomial Logit Modellerinin Tahmini ve Karşılaştırmalı Analizi

Doi: 10.29023/alanyaakademik.1136997

Hilmi KARAALİOĞLU[1]

[1] Öğr. Gör. Dr., Alanya Alaaddin Keykubat Üniversitesi, hilmi.karaalioglu@alanya.edu.tr, Orcid No:0000-0002-1620-5460

***Keywords**:*
Revisit Intention, Parallel Lines, Multinomial Logit, Generalized Ordered Logit, IIA Assumption

Received:
28.06.2022

Accepted:
19.10.2023

ABSTRACT

Although both OLOGIT and MLOGIT models are from the logit model family and their mathematical form is similar, these models have many different aspects besides their similarities. In this context, two dependent variables, one nominal and the other ordinal, measuring the same information were placed in the same data set and these models were compared from various perspectives by making these variables dependent variables. These are the significance of the parameters, their suitability for estimation, ease of implementation, and the provision of assumptions. During the implementation, GOLOGIT was put into practice because Ordered Logit did not provide the parallel regression assumption. Although the number of significant parameters is the same in both models, GOLOGIT stands out in terms of providing detailed analysis for each level of each qualitative independent variable and making fewer model estimations than MLOGIT.

***Anahtar Kelimeler**:*
Yeniden Ziyaret Niyeti, Paralel Regresyon, Multinomial Logit, Genelleştirilmiş Sıralı Logit, IIA Varsayımı

ÖZET

Her ne kadar OLOGIT ve MLOGIT modelleri logit model ailesine ait olsalar ve matematiksel yapıları benzer olsa da bu modeller benzerlikleri dışında birçok farklı yönü de içermektedir. Bu bağlamda, aynı bilgiyi ölçen biri nominal diğeri sıralı iki bağımlı değişken, aynı veri setine yerleştirilmiş ve bu değişkenler bağımlı değişkenler yapılarak bu modeller çeşitli perspektiflerden karşılaştırılmıştır. Bu perspektifler arasında parametrelerin anlamlılığı, tahmini için uygunluk, uygulama kolaylığı ve varsayımların sağlanması yer almaktadır. Uygulama sırasında, Paralel Regresyon Varsayımı'nı sağlamadığı için Sıralı Logit tercih edilemediğinden GOLOGIT uygulamaya konulmuştur. Her iki modelde de anlamlı parametre sayısı aynı olsa da GOLOGIT, her bir kalitatif bağımsız değişkenin her düzeyi için ayrıntılı analiz sağlaması ve MLOGIT'ten daha az model tahmini yapması açısından öne çıkmaktadır.

* Bu makale, yazar Hilmi Karaalioğlu'nun, Prof. Dr. Adil Korkmaz danışmanlığında gerçekleştirilen "Alanya'da Konaklayan Yabancı Turistlerin Esnaf Algılarının Ekonometrik Yöntemlerle İncelenmesi" başlıklı yüksek doktora tez çalışmasından üretilmiş olup, çalışma için Akdeniz Üniversitesi Sosyal ve Beşerî Bilimler Bilimsel Araştırma ve Yayın Etiği Kurulu'nun 29.01.2020 tarihli ve 07 nolu kararınca etik kurul onayı alınmıştır.

1. INTRODUCTION

In regression analysis, it is common to use the Least Squares method when the dependent variable is continuous. However, when the dependent variable is categorical, it is only linear probability models (LPM) that allow the method to be used. These models do not provide the assumptions of linearity, homoscedasticity and autocorrelation, which are the basic assumptions of the ordinary least squares method, and the parameter estimates to be obtained do not fit well with economic realities (Gujarati, 2004). Despite these drawbacks, it is seen that LPM is used in some studies (Genceli, 2011; Kawasaki & Zimmermann, 1964; Uzgören & Uzgören, 2007).

The fact that the dependent variable is categorical is mostly due to the structure of this variable (for example, if the gender is male-female or a machine is faulty-solid). However, in some cases, changing the structure of any continuous variable by the researcher (such as grouping a continuous variable such as income at certain intervals and naming these groups) can put this variable in a categorical structure. In this case, the dependent variable examined in response to the independent variable set in regression models is expressed with the mean value (mean) of the relevant independent variables. The form of this equation is as follows:

$$E(Y|x_i) = b_0 + b_i x_i \qquad i = 1.2. \ldots n \tag{1}$$

where $E(Y|x_i)$, represents the conditional expected value of Y versus the dependent variable set, the regression parameters b_0 and b_i. A notation as in Equation (1) is a method that eliminates all the above-mentioned drawbacks, and the scatter plot of this equation is as follows:

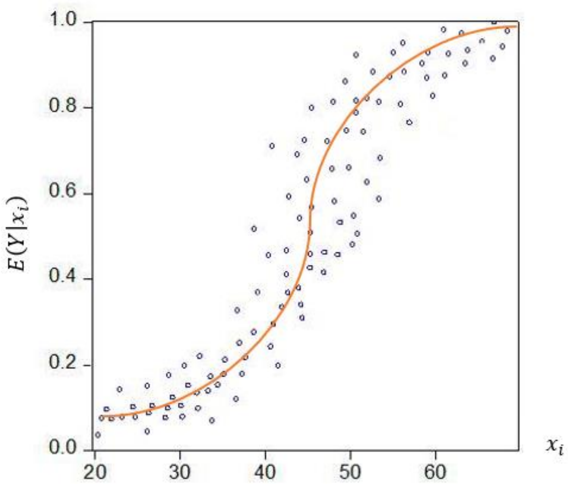

Figure 1. Scatter Plot of $E(Y|x_i)$ according to x_i variables

The functions with the most suitable distribution for the given scatter diagram are the logit distribution and the normal cumulative distribution (probit) functions. Researchers prefer logit or probit models that fit the aforementioned distribution functions in order to avoid the above-mentioned drawbacks of LPM as well as the distribution fits these functions very well. In practice, logit models stand out as more preferred models than probit models due to the fact that logit models have a more flexible structure than probit models, they can be processed mathematically more easily, and estimations are obtained more meaningfully (Karaalioğlu, 2022).

In data sets from which the categorical dependent variable is chosen, the logit model, also known as binary choice models, is a good choice when the dependent variable is two-level. However, the number of levels of the dependent variable mentioned is often more than two. In this case, the methods to be followed differ depending on whether the dependent variable levels are in nominal or ordinal structure. Multinomial or Ordinal Logit models are preferred for dependent variables with categorical levels, provided that they are not missing, reduced or censored (in such cases models such as Tobit, discrete regression models are recommended). In some cases, the researcher can predetermine the nominal or ordinal structure of the dependent variable according to the model he will apply, but this is not possible especially in secondary data sets.

Although both of the mentioned models are from the logit model family, these two models have their own assumptions; In order to provide these assumptions, it becomes necessary to apply some transformations. This is a factor that affects the ease of application and parameter estimation of these models. In addition, there are differences in the interpretations of the obtained estimations due to the scale structure of the dependent variable.

In the literature, it is seen that the assumptions of the ordered logit model are frequently violated (Long, 2006). In this case, researchers hesitate between continuing with the current model or switching to a more complex model to interpret, such as the multinomial model (Williams, 2016). A good alternative to these models is the Generalized Ordered Logit (gologit) models. Gologit models have been used effectively in package software such as Stata with commands such as gologit2 recently. In this study, by looking at the assumptions and goodness of fit of the gologit and mlogit models, it reveals the factors affecting the tourists' intention to revisit Alanya and reveals which one gives more useful results in order to contribute to the sector. In addition to the mentioned purpose, it is desired to compare the gologit and mlogit models using tourism data. Because there has not been a study comparing tourism data and these models as mentioned before. The mentioned comparison is concerned with determining which model yields more successful results in accurately identifying the variables believed to influence tourists' intention to revisit, as well as presenting these variables in a manner consistent with the literature. The aim here is to compare the models with their mentioned aspects and to present them as a guiding tool for future studies.

A significant difference of this study is that it analyzes two different models in the same dataset, using the same independent variables, with two different dependent variables, one nominal and the other ordinal, measuring the same information. The purpose of using the data set of (Karaalioğlu & Korkmaz, 2021) in this study is that this dataset was created by the primary data collection method and contains variables that directly measure the intention of foreign tourists to revisit Alanya. This study consists of five sections. In the first, it is mentioned why logit or probit models are needed when the dependent variable is categorical in a regression model. In the second and third sections, the Multinomial and Ordered logit models, which are the models used in this study, are introduced, respectively. The fourth section includes implementation and analysis, and the concluding section includes discussion and conclusion.

2. MULTINOMIAL LOGIT MODEL / DISCRETE CHOICE MODEL (MLOGIT)

It is quite common to use logit models when the categorical dependent variable is two-level. However, when the number of levels is more than two and these levels are in accordance with the nominal scale, Multinomial logit models, also known as the discrete selection model, should be preferred (McFadden, 1974). In these models, one of the levels of the dependent variable is determined as the baseline and parameter estimates are made so that the other levels give comparative results to this baseline. Such that, in a model where a dependent variable with n levels is examined, the logit model is estimated at the same time as 1 minus the number of levels $(n-1)$. This situation makes it difficult to interpret too many models at the same time, especially in models with a large number of levels, and causes conceptual and notational problems. For the aforementioned reason, there are studies suggesting that the number of levels of the dependent variable should be limited to three levels (Hosmer Jr et al., 2013). To make the structure of the model more understandable, if the output has three levels ($Y_{ref} = 0.\ Y_1 = 1$ and $Y_2 = 2$) with the help of a data set containing k explanatory variables and constant terms, two logit functions $x_0 = 1$ and $\boldsymbol{x}'$ is a vector of length $k+1$. are shown as follows (Hosmer Jr et al., 2013):

$$\mathcal{F}_1(x) = ln\left[\frac{P(Y=1|x)}{P(Y=0|x)}\right] \tag{2}$$

$$= \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \cdots + \beta_{1p}x_k \tag{3}$$

$$= \boldsymbol{x}'\boldsymbol{\beta_1} \tag{4}$$

and

$$\mathcal{F}_2(x) = ln\left[\frac{P(Y=2|x)}{P(Y=0|x)}\right] \tag{5}$$

$$= \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \cdots + \beta_{2p}x_k \tag{6}$$

$$= \boldsymbol{x}'\boldsymbol{\beta_2} \tag{7}$$

The conditional probabilities for the given logit functions are as follows:

$$P(Y=0|x) = \frac{1}{1+e^{\mathcal{F}_1(x)}+e^{\mathcal{F}_2(x)}} \tag{8}$$

$$P(Y=1|x)=\frac{e^{\mathcal{F}_1(x)}}{1+e^{\mathcal{F}_1(x)}+e^{\mathcal{F}_2(x)}} \quad (9)$$

$$P(Y=2|x)=\frac{e^{\mathcal{F}_2(x)}}{1+e^{\mathcal{F}_1(x)}+e^{\mathcal{F}_2(x)}} \quad (10)$$

The aforementioned probability and logit values correspond to three hypothetical levels $(j=0.1.2)$ of the categorical dependent variable. these relations are a function of $2(k+1)$ parameter vectors $\boldsymbol{\beta}'=(\boldsymbol{\beta}'_1,\boldsymbol{\beta}'_2)$. The conditional probability function for a dataset of n volumes with $\pi_j(x)=P(Y=j|x)$ is as follows:

$$l(\beta)=\prod_{i=1}^{n}[\pi_0(x_i)^{y_{0i}}\cdot\pi_1(x_i)^{y_{1i}}\cdot\pi_2(x_i)^{y_{2i}}] \quad (11)$$

Since $\sum y_{ji}=1$ for each i, using the equation (11), the log-likelihood function is obtained as follows:

$$L(\beta)=\sum_{i=1}^{n}y_{1i}\mathcal{F}_1(x_i)+y_{2i}\mathcal{F}_2(x_i)-ln\left(1+e^{\mathcal{F}_1(x_i)}+e^{\mathcal{F}_2(x_i)}\right) \quad (12)$$

The likelihood equations are obtained by taking the partial derivatives of Equation (12) according to each unknown parameter. The general form of these mentioned equations is as follows ($\pi_{ji}=\pi_j(x_i)$ marked for simplicity):

$$\frac{\partial L(\beta)}{\partial\beta_{jk}}=\sum_{i=1}^{n}x_{ki}\,(y_{ji-}\pi_{ji}) \quad (13)$$

The parameter estimations to be obtained for the model, $\widehat{\boldsymbol{\beta}}$, are the roots of the equation (13). Multinomial logit models are essentially a special case of the logit model, and their analysis processes and basic principles are quite similar. Of course, multinomial probit models can be preferred instead of these models. However, the fact that multinomial probit models have some strict assumptions such as the normal distribution of errors, which are not sought in the multinomial logit model, do not make these models very attractive.

In order to obtain the $\widehat{\boldsymbol{\beta}}$ parameters, the representations in the logit model need to be generalized. Since more than one logit model is estimated simultaneously in the multinomial logit model, there should be no multicollinearity problem among the explanatory variables. For the same reason, it is desirable that the number of observations be as large as possible in order to obtain consistent and unbiased estimations. Although the mentioned requirements exist for the model, the assumption of the independence of irrelevant alternatives (IIA) stands out as the most important requirement specific to this model type (Mert, 2016). The mentioned assumption means that if the levels in the categorical dependent variable are increased by one or more levels, the parameter estimation will not change according to its initial state. To put it more clearly, in the binary logit where a categorical dependent variable with levels $\{y_a,y_b\}$ is estimated, the probability of y_a being preferred over the other level should not change in the presence of other levels, according to the IIA assumption. Even if a third level in the form of y_c is added to the model, the probability that y_a will be preferred over level y_b should not change. Otherwise, parameter estimates will become biased and inconsistent as the number of nominal levels changes.

The Hausman-McFadden test developed for the aforementioned assumption compares the parameter estimates of the original model with the constrained model obtained by subtracting one or more of the unrelated levels of the dependent variable. For this, the parameters of the original model obtained after the Multinomial logit model is estimated and the levels of the dependent variable are removed from the model in a controlled manner, and the differences of the parameters of the constrained models to be obtained are tested (Vijverberg, 2011).

Marginal effects can be exploited when interpreting model parameters. Here, the marginal effects do not have to have the same sign as the signs of the parameter estimates as in the logit model. For this reason, risk or difference ratios are considered when explaining marginal effects.

3. ORDERED LOGIT MODEL (OLOGIT)

Multinomial Logit models cannot yield effective results if the categorical dependent variable is an ordinal scale variable, not a nominal one. In this case, the model to be preferred should be Ordered Logit or Ordered Probit models. In that case, the basis of Ordered Logit models is that the dependent variable is a categorical, more than two-level and ordinal variable.

In the mentioned model type, no additional assumptions are sought for the independent variables other than the classical linear regression assumptions, as in MLOGIT. However, due to the categorical nature of the dependent variable, normality and homoscedasticity conditions will not be sought in these models. However, unlike MLOGIT, since a single model is estimated for all dependent variable levels, the 'parallel regression assumption' stands out as an important requirement that the model must provide. This requirement, which is called 'proportional odds assumption' or 'proportional odds assumption' in various sources, means that for each term added to the model, the parameter estimates obtained for the binary levels of the dependent variable do not change depending on which level the analysis is made (Quanticate, 2016. Jun 10). Accordingly, the effects of any explanatory variable should be consistent or proportional at different dependent variable levels. For this reason, the relevant requirement is also called the 'proportional probability assumption'. With the help of the test developed by Brant, it can be checked whether it has been violated (Brant, 1990). If the requirement mentioned in the OLOGIT model is found to be violated, the parameter estimates of the model may lead to inaccurate or misleading results. Therefore, Generalized Ordered Logit (GOLOGIT), which allows relaxation of the parallel regression assumption, should be preferred instead of OLOGIT. Because in the estimation of GOLOGIT, the coefficients do not have to have equal probability between the levels of the independent variable. In the absence of theoretical or different a priori information about which variable may have a random effect, the researcher should prefer to model with as few variables as possible. In this way, explanatory variables that have strong effects on the prediction can be modeled with random effects if the variances of these effects are significant enough as evidenced by their significance and magnitude (Snijders, 2005). However, care should be taken not to increase the number of variables with random effects too much and not to transform the model into a cumbersome one.

The generalized ordered logit model for a dependent variable with m levels can be represented as (Williams, 2016):

$$P(Y_i > j) = F(\beta_j x_i) = \frac{e^{(\alpha_j + \beta_j x_i)}}{1+e^{(\alpha_j + \beta_j x_i)}} \qquad j = 1.2. \dots m-1. \tag{14}$$

Assuming that a variable has M levels, the coding for the level where the preference of any unit hits is commonly continued as 1 for the first level, 2 for the second level, and finally to m for the $M-th$ level. Thus, a logical order is created for the dependent variable levels. The point to be noted here is that this ordering is a logical order, and the values do not contain any meaning other than semantic ordering. Therefore, Y_i values that only express order will not give logical results when subjected to algebraic operations.

Equation (14) more clearly is as follows:

$$P(Y_i = 1) = 1 - \frac{e^{(\alpha_1 + \beta_1 x_i)}}{1 + e^{(\alpha_1 + \beta_1 x_i)}} \tag{15}$$

$$P(Y_i = j) = \frac{e^{(\alpha_{j-1} + \beta_{j-1} x_i)}}{1 + e^{(\alpha_{j-1} + \beta_{j-1} x_i)}} - \frac{e^{(\alpha_j + \beta_j x_i)}}{1 + e^{(\alpha_j + \beta_j x_i)}} \tag{16}$$

$$P(Y_i = M) = \frac{e^{(\alpha_{m-1} + \beta_{m-1} x_i)}}{1 + e^{(\alpha_{m-1} + \beta_{m-1} x_i)}} \tag{17}$$

The maximum likelihood method is preferred for parameter estimations of the model. The odds ratios of the cumulative frequencies to be obtained in this way are as follows(Ugurlu, 2015):

$$P(Y_i \leq m) = \sum_{j=1}^{m} P(Y = j) \tag{18}$$

The odds ratio to be obtained in the model estimation is expressed as the difference between those greater than any m level of a level and those less than m equal. More precisely, the ratio mentioned will be the difference between the probability $P(Y_i > m)$ and the probability $P(Y_i \leq m)$ and is represented as follows:

$$R_m = \frac{P(Y_i \leq m)}{P(Y_i > m)} = \frac{P(Y_i \leq m)}{1 - P(Y_i \leq m)} \tag{19}$$

In the given equation, odds ratios are independent of the number of levels and odds ratios are assumed to be constant for all levels. In this case, the difference of the sequentially given levels with the hierarchically lower levels is calculated. The e^{β_i} value to be obtained as a result is calculated as the odds ratio.

The equation given in Equation (14) can be generalized under m ordered levels as follows:

$$\begin{gathered} P_m = F(\boldsymbol{\beta}'\boldsymbol{x}) \\ P_{m-1} = F(\boldsymbol{\beta}'\boldsymbol{x} + \alpha_1) - F(\boldsymbol{\beta}'\boldsymbol{x}) \\ P_{m-2} = F(\boldsymbol{\beta}'\boldsymbol{x} + \alpha_1 + \alpha_2) - F(\boldsymbol{\beta}'\boldsymbol{x} + \alpha_1) \\ \vdots \end{gathered} \tag{20}$$

These equations will also provide the following probability sums:

$$\begin{aligned} & P_m = F(\boldsymbol{\beta}'\boldsymbol{x}) \\ & P_m + P_{m-1} = F(\boldsymbol{\beta}'\boldsymbol{x} + \alpha_1) \\ & P_m + P_{m-1} + P_{m-2} = F(\boldsymbol{\beta}'\boldsymbol{x} + \alpha_1 + \alpha_2) \\ & \qquad \vdots \\ & P_m + P_{m-1} + \cdots + P_2 = F(\boldsymbol{\beta}'\boldsymbol{x} + \alpha_1 + \alpha_2 + \cdots + \alpha_{m-2}) \\ & P_1 = 1 - F(\boldsymbol{\beta}'\boldsymbol{x} + \alpha_1 + \alpha_2 + \cdots + \alpha_{m-2}) \end{aligned} \tag{21}$$

In practice, the $\alpha_1 + \alpha_2 + \cdots + \alpha_{m-2}$ parameters in equations (21) are not required to be positive. Maximum likelihood estimates will generally give positive estimates for these parameters. Otherwise, there may be some specification errors in the model (Maddala, 1983).

As stated in Equations (20) and (21), the probability for any level m can be shown as $P_m = F(\boldsymbol{\beta}'\boldsymbol{x})$. The probability can then be expressed as:

$$P(Y_i = m|x_i) = F(x_i\beta_m + \alpha_m) - F(x_i\beta_{m-1} + \alpha_{m-1}) \tag{22}$$

The likelihood function for determining the model parameters is as follows:

$$\begin{aligned} L(\beta_j, \alpha_j | Y_i, x_i) &= \prod_{j=1}^{m} \prod_{Y_i = j} P(Y_i = j | x_i, \beta_j, \alpha_j) \\ &= \prod_{j=1}^{m} \prod_{Y_i = j} F(x_i\beta_j + \alpha_j) - F(x_i\beta_{j-1} + \alpha_{j-1}) \end{aligned} \tag{23}$$

When the logarithm of the equation is taken, the likelihood function is obtained as follows:

$$L(\beta_j, \alpha_j | Y_i, x_i) = \sum_{j=1}^{m} \sum_{Y_i = j} F(x_i\beta_j + \alpha_j) - F(x_i\beta_{j-1} + \alpha_{j-1}) \tag{24}$$

The parameters of the model are determined by maximizing the logarithmic likelihood function with the help of derivative. The resulting predictors are asymptotically normal, efficient, and consistent (Ugurlu, 2015).

4. IMPLEMENTATION AND ANALYSIS

In this study, the survey data compiled by Karaalioğlu (2022) for the purpose of "Analysis of Tradesmen Perceptions of Foreign Tourists Visiting Alanya Using Econometric Methods" were used. As previously stated, the reason for using the mentioned dataset is its status as a "primary dataset" that examines the impact of tourism businesses in Alanya on foreign tourists' intention to revisit, with the aim of making purposeful inferences. In this concept, Multinomial and Ordered Logit models were created, with two dependent variables, respectively, one nominal and the other ordinal, that measure almost the same information in the aforementioned data set.

The first of the mentioned dependent variables is the variable 'I will come to Alanya again for a holiday', which measures the tourist's revisit intention. The levels of this variable are "I will come", "I will not come" and "I have no idea", and the level of "I will not come" is arranged as the control group. A result variable with any number of levels can be used to represent the scope of the model and methods. But the details are most easily illustrated by three categories. Generalizing further to more than three categories is a problem of representation rather than concept (Hosmer, 2000). In this context, the levels of "strongly disagree" and "disagree" are defined as "I will not come" since they are expressions that express negative judgment. The levels of 'strongly agree' and 'agree' are categorized as 'my future' because they were expressions expressing positive judgment. Another purpose of combining the categories is that the OLOGIT model is used in the study. The aim of this study is to compare MLOGIT and OLGIT (vs. GOLOGIT) models from various perspectives. In the OLOGIT model, the cells of the cross-tables obtained between the dependent variable and the independent qualitative variables should be sufficiently filled. In other words, there should be no zero or empty cells. Because empty cells cause the problem that the test statistics of the Brant test, which is used for testing the parallel regression assumption, which is an important assumption of OLOGIT, cannot be calculated (Mert, 2016). For this reason, the levels of the dependent variable in the OLOGIT model must be combined.

The levels of the MLOGIT model are also combined in order to have a standard similarity with the dependent variable of the OLGOIT model and to avoid conceptual or notational problems and to provide a simpler interpretation of the parameters economically. The mentioned levels are combined by renaming as mentioned above. Table 4 shows the status of dependent variable levels (and some independent variable levels) before combining; In the table, empty or zero cells can be seen.

The dependent variable of the model OLOGIT is 'The effect of local shopkeepers and shopping experience on the intention to revisit Alanya'. The levels of this variable were categorized as 'very negative', 'negative', 'moderate', 'positive' and 'very positive' in terms of their impact on revisit intention. If the levels of a categorical variable can be ordered from negative to positive or from negative to positive, it is a sortable qualitative variable. As the variable is in ordinal form, it is accepted as the dependent variable of the Ordered Logit model. In this model, the dependent variable levels are categorized as 'negative idea', 'no idea (indecisive)' and 'positive idea' for the reasons mentioned both above and in the continuation of the text. While the codes of the aforementioned levels were in order from 0 to 5. the first two levels were 0. the moderate level 1 and the last two levels 2. and the level reduction was carried out on the data set.

The variables used as independent variables for both models and the coding of these variables are as follows: Gender (M:1. F:0), working status (working in a paid job (salary): 1. Retired :2. Own business (employer, etc.) .):3. Unemployed (looking for a job):4. Housewife, etc.:5. Student:6. other:7), income level by country of residence (Very low:1. low:2. middle:3. high :4. very high:5), the ratings of Alanya shopkeepers' attitudes (very bad:1. bad:2. medium:3. good:4. very good:5) and reduced factors obtained by factor analysis of variables in the context of tourists' perception of tradesmen. The factors mentioned are the variables coded as F1 - F5 and can be seen in Table 3.

Since more than one model is estimated simultaneously in MLOGIT and the number of independent variables is high, before starting the analysis, the requirement that the number of observations, which is a necessary assumption for the model, should be sufficiently large. In this context, it was seen that 446 people out of 564 gave answers as 'I will come to Alanya again', 88 people 'I have no idea' and 30 people 'I will not come'. Since there are no small frequencies in the frequency table, the assumption that there are no small-frequency cells required for the model is provided.

Table 1. Model Fit Information for Multinomial Logit Model

Model Fit Information			
		Chi-square	***p***
Likelihood Ratio Test		196.638	0.000
Goodness of Fit	Pearson	805.852	0.000
	Deviance	456.715	0.992

In the table, first of all, the significance of the estimated model is examined by looking at the p value and the likelihood ratio chi-square values. Here, the model is significant since $chi - square = 196.638$ and $p = 0.00 < 0.05$. Deviance and Pearson chi-square tests should also be examined, which are useful in determining whether a model fits the data well. Accordingly, non-significant test results ($p > 0.05$) are indicators that the model fits the data well.

While estimating the model, the independence of irrelevant alternatives (IIA) assumption needs to be examined. In addition, the Hausman test result for testing this assumption is shown in Table 3. According to the Hausman test specified in the table, the IIA assumption is valid, and the model is applicable. Another test of this assumption is as follows: the levels of the predicted model and the dependent variable are removed one by one from the model, and the parameters of the models obtained are compared and the difference between these parameters is tested. In the model, 'I have no idea' and 'I will come' levels other than the baseline level of 'I will not come' of the dependent variable should be removed from the model one by one and it should be checked whether the parameter estimates have changed significantly. In order to provide the IIA assumption, the differences between the parameters of all the predicted models should be tested and these differences should be statistically insignificant. Otherwise, the IIA assumption is deemed to have been violated, since significant differences would mean significantly different parameters from each other.

Table 3 shows the parameter estimates of the Multinomial logit model. As it is known, regression estimation should be made in the model by one less than the number of levels. Here, two regression estimates are shown against three levels of the dependent variable. It is shown in the table that the category 'I will not come', numbered with '0', is the reference category. The estimated regression equations are those obtained relative to this reference category. In this context, the models obtained for the 'I have no idea' and 'I will come' levels of the variable 'YZN: I will come to Alanya again for a holiday', which measures the foreign tourist's revisit intention, were compared with the reference category (control group) 'I will not come' level (either are also relatively obtained models.).

As for the estimation of the Ordered logit model, there are no assumptions that the independent variables must provide, apart from the accuracy of the mathematical form in the model and the requirement to ensure multicollinearity for the continuous variables. Therefore, explanatory variables can take any form.

Table 2. IIA Test Results for the Multinomial Logit Model

```
. test [full_1=gel_1], cons

(1) [full_1]Factor1 - [gel_1]Factor1=0
(2) [full_1]Factor2 - [gel_1]Factor2=0
(3) [full_1]Factor3 - [gel_1]Factor3=0
(4) [full_1]Factor4 - [gel_1]Factor4=0
(5) [full_1]Factor5 - [gel_1]Factor5=0
(6) [full_1]Gender - [gel_1]Gender=0
(7) [full_1]Workingstatus - [gel_1]Workingstatus=0
(8) [full_1]Incomelevel - [gel_1]Incomelevel =0
(9) [full_1]Tradesmenrate - [gel_1]Tradesmenrate=0
(10) [full_1]_cons - [gel_1]_cons=0

        Chi2(10)        = 19.09
        Prob > Chi2     =  0.392

---------------------------------------------------
. test [full_2=kar_2], cons

(1) [full_2]Factor1 - [kar_2]Factor1=0
(2) [full_2]Factor2 - [kar_2]Factor2=0
(3) [full_2]Factor3 - [kar_2]Factor3=0
(4) [full_2]Factor4 - [kar_2]Factor4=0
(5) [full_2]Factor5 - [kar_2]Factor5=0
(6) [full_2]Gender - [kar_2]Gender=0
(7) [full_2]Workingstatus - [kar_2]Workingstatus=0
(8) [full_2]Incomelevel  - [kar_2]Incomelevel =0
(9) [full_2]Tradesmenrate - [kar_2]Tradesmenrate=0
(10) [full_2]_cons - [kar_2]_cons=0

        Chi2(10)        = 12.62
        Prob > Chi2     =  0.246
```

In addition, normality and homoscedasticity assumptions are not examined in these models due to the structure of the dependent variable. However, the 'parallel regression assumption' in Ordered logit is an important assumption to examine. This assumption gains importance because a single equation is estimated for all levels of the dependent variable in the Ordered logit model. This assumption implies that the pass-through has an equal probability for all levels of the dependent variable.

Because the dependent variable is an ordinal-scale variable, there are intervals that are considered to be logically equal between the ordinal levels. If the parallel regression assumption is not met, the linear estimators will not be the same at each level. In this case, Generalized Ordered logit models should be used, which require estimation separately for each level. For this assumption, the frequencies of the cross-tables between each level of the dependent variable and the qualitative independent variables should be at a sufficient level (or not empty). For this reason, if we look at Table 4 given as an example, zero or low frequency cells stand out.

Table 3. Parameter Estimates for the Multinomial Logit Model

Levels / Independents	Revisit intention: no idea*				Revisit intention: positive*			
	β	p	exp (β)	Odds ratio	β	p	exp (β)	Odds ratio
constant	-4.265	**0.037**	-2.09	0.014	-6.999	**0.000**	-3.52	0.001
F1: Perception of product presentation	0.981	**0.000**	4.45	2.669	1.510	**0.000**	6.79	4.530
F2: The attitude of the tourist towards the cooperation of the tradesmen with the commissioners.	-0.404	**0.003**	-2.97	0.667	-0.532	**0.000**	-4.03	0.587
F3: Dissatisfaction with post-purchase shopping.	-0.209	0.193	-1.30	0.811	-0.308	**0.049**	-1.97	0.734
F4: Attitude of the tourist towards hotel services.	0.687	**0.005**	2.80	1.988	0.610	**0.011**	2.54	1.840
F5: Attitude of the tourist towards product prices	-0.765	**0.024**	2.26	0.465	-0.731	**0.027**	2.21	0.481
Gender[1]	-1.578	**0.003**	-2.94	0.206	-1.746	**0.001**	-3.35	0.174
Working status[2]	-0.041	0.808	-0.24	0.959	-0.277	0.099	-1.65	0.757
Income level[3]	0.498	0.121	1.55	1.645	0.420	0.173	1.36	1.522
Tradesmen rating	1.188	**0.000**	-3.72	3.282	0.926	**0.003**	-2.95	2.526
Hausman Test: $\mathcal{X}^2(18) = 0.69$ $Prob > chi2 = 1.000$								
Reference categories: *negative, [1]female, [2]working in a paid job (salary), [3]very low								

Since zero or low frequency cells are seen as a problem in terms of providing the parallel regression assumption, row or column combinations should be used in these tables. In this context, it is necessary to combine some levels in both the dependent variable and the independent variable. This inevitably causes the structure of the variables and therefore the data set to change.

Table 4. Cross Frequency Table between Dependent Variable (SYZN) and Working Status

Dependent Variable	*Working Status*						
	Working in a paid job (salary)	**Retired**	**Own business (employer, etc.)**	**Unemployed (looking for work)**	**Housewife, etc.**	**Student**	**Total**
Very negative	58	6	6	4	2	0	**76**
Negative	60	4	12	4	8	0	**88**
No idea	130	12	20	4	8	8	**182**
Positive	82	12	12	4	6	4	**120**
Very positive	66	10	12	0	8	2	**98**
Total	**396**	**44**	**62**	**16**	**32**	**14**	**564**

Table 5. Likelihood Ratio Test Results for Testing the Parallel Regression Assumption in OLOGIT

The effect of tourists' interactions with local shopkeepers and their shopping experiences on their intention to revisit Alanya (SYZN)	β	p	exp(β)
F1: Perception of product presentation	1.544	**0.000***	14.61
F2: The attitude of the tourist towards the cooperation of the tradesmen with the commissioners.	0.109	0.056	1.91
F3: Dissatisfaction with post-purchase shopping.	0.013	0.836	0.21
F4: Attitude of the tourist towards hotel services.	-0.169	**0.065****	-1.84
F5: Attitude of the tourist towards product prices	0.084	0.489	0.69

Gender[1]	Male	-0.508	**0.018**	-2.37
Working status[2]	Retired	-0.685	**0.074****	-1.79
	Own business (employer, etc.)	0.184	0.584	0.55
	Unemployed (looking for work)	-1.351	**0.047***	-1.99
	Housewife, student, etc.	0.212	0.614	0.50
Income level[3]	Middle - income level	-1.106	**0.000***	-3.58
	High - income level	-0.247	0.516	-0.65
Tradesmen rating[4]	Average	-0.471	0.321	-0.99
	Decent	-0.727	0.140	-1.47
Cut 1		13.591		
Cut 2		16.815		
$Pseudo\ R2\ = 0.4593$ $Prob > chi2\ = 0.0000$ $LR\ Chi2\ (14)\ = 565.52$				
Reference categories: [1]female, [2] working in a paid job (salary), [3]low-income level ,[4]bad				

After the aforementioned level combinations, dummy variables were created and the results of the Likelihood Ratio Test to be used in the test of the parallel regression assumption are given in Table 5. The most important part of these results is the $Prob > chi2 = 0.000$ result at the bottom. Here, since $Prob = 0.000 < 0.05$, the null hypothesis, which states that the odds ratios do not change between the categories of the SYZN dependent variable, will be rejected. In other words, the assumption of parallel regression will not be provided.

There are too many parameter estimations in the generalized ordered logit analysis, as in the mlogit model. This is because the parallel regression assumption is broken. The parallel regression assumption may be corrupted by several independent variables in the model or may be corrupted for only one independent variable. Thus, due to one or more variables, all parameters are estimated over and over again. In the generalized ordered logit analysis, it is possible to determine the independent variables that cause the parallel regression assumption to not be met, and to make separate parameter estimation for these variables according to the regimes. In this case, the same parameters can be used in all regimes for other independent variables. To see this, a comparison of the gologit2 command was made at the 5% significance level. In this comparison, just as in the Brant test, the independent variables that cause the parallel regression assumption not to be met are determined. While different parameter estimations are made for these variables (for each regime), the same parameter estimations are made for other variables. Test results are given in Table 6:

Table 6. Testing Parameter Lines Assumption of the Generalized Ordered Logit Model

```
.gologit2 average tradesmen rating decent tradesmen rating  middle-income level  high-income level  retired own
business (employer, etc.) unemployed housewife(student, etc.) male F1 F2 F3 F4 F5

Testing parallel line assumption using the .05 level of significance…
Step 1:   Constraints for parallel lines imposed for middle-income level (P value = 0.8704)
Step 2:   Constraints for parallel lines imposed for high-income level (P value = 0.9342)
Step 3:   Constraints for parallel lines imposed for own business (employer, etc.) (P value = 0.7810)
Step 4:   Constraints for parallel lines imposed for F1 (P value = 0.6382)
Step 5:   Constraints for parallel lines imposed for F2 (P value = 0.5799)
Step 6:   Constraints for parallel lines imposed for F3 (P value = 0.4723)
Step 7:   Constraints for parallel lines imposed for F5 (P value = 0.0741)
          F4 (P value = 0.00053)
          male (P value = 0.00000)
          retired (P value = 0.00046)
          unemployed (P value = 0.00000)
          housewife (student, etc.) (P value = 0.03702)
          average tradesmen rating (P value = 0.00000)
          high tradesmen rating (P value = 0.00004)

Wald test of parallel lines assumption for the final model:
    (1)     [1] middle-income level – [2] middle-income level=0
    (2)     [1] high-income level - [2] high-income level=0
    (3)     [1] own business (employer, etc.)  - [2] own business (employer, etc.) =0
    (4)     [1] F1 - [2] F1=0
    (5)     [1] F2 - [2] F2=0
    (6)     [1] F3- [2] F3 =0
    (7)     [1] F5- [2] F5 =0


            chi2(6)       =  4.28
            Prob > chi2   =  0.7472
```

An insignificant test statistic indicates that the final model
does not violate the proportional odds / parallel lines assumption

If you re-estimate this exact same model with gologit2. instead
of autofit you can save time by using the parameter

p1 (middle-income level high-income level own business (employer, etc.) F1 F2 F3 F5)

Reference categories: [1]female, [2] working in a paid job (salary), [3]low-income level ,[4]bad ,[5]Negative ,[6]Negative + No idea (indecisive)

The results given in Table 6 show whether the parameters estimated for each independent variable show a significant difference between the regimens (from one level of the dependent variable to another level). In order for these tests to be performed successfully in statistical software (e.g., STATA) and for the model estimation to be effective and consistent, dummy variables should be created for qualitative variables. According to the results, for example, the parameter of the 'middle-income level' level does not show a significant difference at the 0.05 significance level from the level of the dependent variable (p=0.8704>0.05). Therefore, the dummy variables created for the income level variable, the 'own business (employer, etc.)' variables and the F1, F2, F3, F5 factors are not the variables that cause the parallel regression assumption to be broken. In this case, the assumption in question can be imposed on these variables; A parsimonious model estimation can be performed without extra coefficient estimation in the model. However, since it completely excluded the F4 factor and gender variable, many levels of the working status variable and the tradesmen rating variable, the effects of which were especially desired to be observed within the scope of the study, instead of using a parsimonious model, the GOLOGIT model was put into practice in a way to include all the independent variables in question; In addition, it is aimed to bring a standard interpretation by using all the variables used in the MLOGIT model here as well. By determining the variables that break the parallel regression assumption, a separate model estimation can be made only for these variables. Therefore, although the Generalized ordered logit model is similar to the Multinomial logit model, it is superior to the Multinomial logit model in this respect since less parameter estimation is made. This is one of the reasons for choosing the GOLOGIT model.

Table 7. Parameter Estimates of the Generalized Ordered Logit Model

The effect of tourists' interactions with local shopkeepers and their shopping experiences on their intention to revisit Alanya (SYZN)		**No idea (indecisive) + Positive[5]**				**Positive[6]**			
		$\boldsymbol{\beta}$	$\boldsymbol{p}$	$\exp(\boldsymbol{\beta})$	**Odds ratio**	$\boldsymbol{\beta}$	$\boldsymbol{p}$	$\exp(\boldsymbol{\beta})$	**Odds ratio**
Constant		-18.85	**0.000**	-11.58	6.450	-16.49	**0.000**	-11.23	6.870
F1: Perception of product presentation		1.737	**0.000**	13.86	5.682	1.737	**0.000**	13.86	5.682
F2: The attitude of the tourist towards the cooperation of the tradesmen with the commissioners.		0.093	0.124	1.54	1.097	0.093	0.124	1.54	1.097
F3: Dissatisfaction with post-purchase shopping.		0.012	0.858	0.18	1.012	0.012	0.858	0.18	1.012
F4: Attitude of the tourist towards hotel services.		0.093	0.506	0.67	1.097	0.528	**0.000**	4.00	1.589
F5: Attitude of the tourist towards product prices		0.084	0.518	0.65	1.088	0.084	0.518	0.65	1.088
Gender[1]	Male	0.548	0.098	1.66	1.729	-1.54	**0.000**	-5.01	0.213
Working status[2]	Retired	-2.35	**0.000**	-3.96	0.094	0.367	0.526	0.63	1.443
	Own business (employer, etc.)	0.501	0.164	1.39	1.651	0.501	0.164	1.39	1.651
	Unemployed (looking for work)	-2.70	**0.001**	-3.21	0.066	-1.89	**0.044**	1.01	0.641
	Housewife, student, etc.	1.828	0.092	1.69	6.225	-0.54	0.255	-1.14	0.579
Income level[3]	Middle - income level	-0.94	0.004	-2.88	0.387	-0.94	**0.004**	-2.88	0.387
	High - income level	-0.41	0.314	-1.01	0.663	-0.41	0.314	-1.01	0.663
Tradesmen[4] rating	Average	2.385	**0.001**	3.38	10.86	2.713	**0.000**	4.07	10.66
	Decent	1.307	0.073	1.79	3.696	1.698	**0.008**	2.65	5.228
$Pseudo\ R2\ = 0.5237$ $Prob > chi2\ = 0.0000$ $LR\ Chi2\ (14)\ = 644.78$									
Reference categories: [1]female, [2] working in a paid job (salary), [3]low-income level ,[4]bad ,[5]Negative ,[6]Negative + No idea (indecisive)									

In the generalized ordered logit model, the probabilities between the dependent variable levels do not need to be equal. The advantage of this model is that it does not impose the parallel regression assumption, which is a major limitation of the ordered logit model. In this model, the coefficients of the variable may differ between the categories of the variables. Therefore, a separate parameter estimation is made for each regime of the dependent variable. In this study, the SYZN dependent variable has 5 levels. However, as stated in the previous section, the number of levels was reduced to 3 by combining rows or columns. In order to ensure simplicity in model estimation, analysis was carried out using 3 levels. For generalized ordered logit estimation, as in the above section, dummy variables should be created for qualitative independent variables.

In summary, the reason for choosing the generalized ordered logit model is that the parallel regression assumption could not be met. This assumption may be due to one or more independent variables in the model or one or more of the dummy variables created for parameter estimation. Therefore, by determining the variables that cause the parallel regression assumption not to be met, a separate model estimation can be made only for these variables. Therefore, although the Generalized Ordered logit model is similar to the Multinomial logit model, it can be said that it is a superior model to the Multinomial logit in this respect since fewer parameters are estimated.

In this context, in order to determine the variables that cause the parallel regression assumption not to be met in the model, all parameters are compared at the 5% significance level. Thus, it is tested whether the parameters of each of the independent variables vary between models. The parameter estimates of the variables that do not violate the assumption will remain the same in each model, while the parameter estimates of the variables that do not violate the assumption will change as the model differs.

After the parameter estimation of the generalized ordered logit model is made, the relative odds ratios in the final model can be obtained from this model. The parameter estimates of the variables that do not satisfy the parallel regression assumption remain the same in both regimes, and the relative odds ratios are also the same. However, the relative odds ratios of the variables, which cause the parallel regression assumption not to be met, differ according to the models, such as their parameters.

Table 8. AIC and BIC Values for MLOGIT, OLOGIT and GOLOIT Models

Model	Akaike Information Criterion (AIC)	Bayesian Information Criterion (BIC)	*p*
Multinomial Logistic Regression	716.314	725.044	.0000
Ordered Logistic Regression	712.024	719.709	.0000
Generalized Ordered Logistic Regression	704.624	711.325	.0000

As a result, the models put into practice within the scope of the study were evaluated in the context of the model selection criterion (goodness of fit). Table 8 shows that Akaike (AIC) and Bayesian (BIC) information criterion are significant for all three models. Considering that the same data set and the same independent variables are preferred for the three models; For the SYZN dependent variable in ordinal structure, the GOLOGIT model is more suitable for the data set than the OLOGIT model. It is seen that the MLOGIT model applied for the nominal dependent variable YZN is also a significant model.

5. DISCUSSION AND CONCLUSIONS

The first of the models applied within the scope of the study is the Multinomial logistic regression model. It is shown that the independence of irrelevant alternatives (IIA) assumption is satisfied before parameter estimation is made. As a result of statistical tests, two variables were not statistically significant; On the other hand, it has been seen that all of the factors obtained as a result of the factor analysis, except the working status of the tourists and the income level according to the countries they live in, are significant. According to these results, the increase in satisfaction with the quality of local products and the satisfaction of foreign visitors with the behavior exhibited by the tradesmen strengthen the intention of these people to visit Alanya again.

On the other hand, awareness of the cooperation between commissioners and local tradesmen reveals itself as a factor that negatively affects the tourist's intention to revisit. Situations such as realizing that there is gender discrimination during shopping, developing suspicions against the seller eventually after shopping, thus falling into distrust, realizing that the label prices are exaggerated, cause foreign visitors to be alienated from shopping, local tradesmen, and commissioners. As a result, this situation emerges as a phenomenon that negatively affects their tendency to come to Alanya for the second time.

Another factor affecting the tourist's intention to revisit is the attitude towards hotel services. If this attitude is positive, revisit intention is strengthened, as expected. Thus, the probability of the 'I will not come' level rises by regressing the probabilities of the other levels. The attitude towards product prices also has an impact on revisit intention. This proves that price competition does not disappear completely but continues to a certain extent. In this context, while the satisfaction of the price levels strengthens the tourist's intention to visit again, it significantly increases the probability of the "I will not come" level compared to other levels. This is true for all genders, male or female; however, the level of 'I will come' for males and 'I have no idea' for females has a greater share of this increase.

Another model applied within the scope of the study is the Ordered logit regression model. For the model, the revisit intentions of the tourists were analyzed by scaling the levels ordinally with the variable 'The effect of the local tradesmen and shopping experience on the intention to revisit Alanya' in a way that interacts with the local tradesmen and shopping experiences. In order to provide the parallel regression assumption, which is one of the important assumptions of the model, low-frequency cells were merged. Although this process provides convenience in the application and interpretation of the parameters, it causes loss of level and reduction of detail. It was seen that the parallel regression assumption was not met despite the merging of the levels and the data set was not suitable for the Ordered Logit model. For this reason, the analysis continued with the Generalized Ordered Logit model instead of ordered logit. The positive aspect of this model compared to Ordered Logit is that it does not require the assumption of parallel regression. In the model, two regimes were determined for three levels of the dependent variable and parameter estimation was made according to these regimes. While the first regime corresponds to the 'negative' level of the dependent variable; the second regime corresponds to the sequential combination of 'negative + no idea(indecisive)' levels. In the multinomial logit, each parameter interpretation is made according to the base level, while in this model, the interpretations are made according to the levels corresponding to the regimes.

In the application of the Generalized Ordered Logit model, dummy variables were created for each level of the qualitative independent variables. This has revealed a model in which the number of independent variables appears to be high. However, this gives an opportunity for detailed analysis for each level of the independent variable. So much so that a basic level is formed within each variable and the effect of the difference between other levels and this level on the dependent variable can be observed. In addition, as a result of such an approach, it can be seen that the parallel regression assumption is not provided by the dummy variable corresponding to which level. Therefore, a model estimation was also made for the variables that did not provide with the assumption in the model. In the Generalized Ordinal logit model, since two different models are estimated for the variables that provide the parallel regression assumption and those that do not, although this model is similar to the Multinomial logit model in application, it provides much less parameter estimation than a multinomial logit model in which the number of levels of the dependent variable will be higher. In this respect, it stands out that it is a superior model than the Multinomial logit model.

In the Generalized Ordered Logit model, among the factors that were statistically significant in the other model, the attitude of the tradesmen towards cooperation with the commissioners, the dissatisfaction with the shopping and the attitude of the tourist towards the hotel services were not found statistically significant. However, in this model, dummy variables corresponding to the 'retired' and 'unemployed' levels of the working status variable and the dummy variables corresponding to the 'middle income group' of the income level variable, which are not significant in the Multinomial logit model, were found to be significant.

Model fits for each of the preferred models are statistically significant. In the context of model selection criteria, the model that is most suitable for the data set and has the highest relative quality is the GOLOGIT model. In addition to the above-mentioned positive aspects of the aforementioned model, one more plus can be said in terms of its suitability for the data set. It is seen that the MLOGIT model, which has a different dependent variable structure, is also a statistically significant model. The fact that the obtained parameters are different in different models is due to both the choice of the model, the detailing of the levels in the GOLOGIT model with the help of dummy variables, and the fact that the dependent variables are different variables (even though they measure very similar information in the same data set).

In conclusion, one of the words that can be said is as follows: Since foreign tourists live in front of local tradesmen, their acquisitions have important effects on their intention to visit Alanya again. In this case, it can be said that local tradesmen, commissioners, and guides should be made aware of not being deceived by foreign tourists. This awareness can also be expressed as follows: Although deception brings a short-term gain, it leads to a complete loss of respectability, alienates foreign visitors from shopping and thus undermines the long-term high incomes of local tradesmen and commissioners; however, not cheating has a completely different result, first of all, it gains prestige, ensures trust in local tradesmen and commissioners, and thus opens the door to a long-term middle income. The important thing in economic life is not to make big profits for once, but to make small, small, or medium-sized gains unlimited times. This, of course, is a phrase that normative economics can say, not positive economics.

REFERENCES

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, *46*(4), 1171-1178.

Genceli, M. (2011). Nitel bağımlı değişken için doğrusal olasılık modeli. *İstanbul Üniversitesi İktisat Fakültesi Mecmuası*, *45*(1-4).

Gujarati, D.N. (2004). *Basic econometrics* (Fourth Edition ed.). The McGraw−Hill Companies.

Hosmer Jr, D.W., Lemeshow, S., & Sturdivant, R.X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.

Hosmer, D.W., & Lemeshow, S. (2000). *Applied logistic regression* (Second edition ed.). Wiley New York.

Karaalioğlu, H. (2022). *Alanya'da konaklayan yabancı turistlerin esnaf algılarının ekonometrik yöntemlerle incelenmesi* (Publication Number EKO - 48) [Doktora Tezi, Akdeniz Üniversitesi].

Karaalioğlu, H., & Korkmaz, A. (2021). Poisson regression analysis with tourism data: analysis of the effects of foreign visitors' local tradership and shopping perceptions on the frequency of visit. *Alanya Akademik Bakış*, *5*(3), 1443-1456.

Kawasaki, S., & Zimmermann, K.F. (1964). Measuring relationships in the log-linear probability model by some compact measures of association. *Statistische Hefte*, *22*(2), 82-109.

Long, J.S., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata* (Vol. 7). Stata press.

Maddala, G.S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511810176

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, *3*(4), 303-328.

Mert, M. (2016). *SPSS STATA yatay kesit veri analizi bilgisayar uygulamaları.* Ankara: Detay Yayıncılık.

Quanticate, t. C.D.E. (2016. Jun 10). *Understanding the proportional odds assumption in clinical trials*. https://www.quanticate.com/blog/understanding-the-proportional-odds-assumption-in-clinical-trials

Snijders, T.A. (2005). Fixed and random effects. *Encyclopedia of statistics in behavioral science*, *2*. 664-665.

Ugurlu, E. (2015). *Kesikli Seçim Modelleri (Logit, Multinomial Logit, Ordered Logit, Sıralı Lojit, Lojistik Model, Marjinal Etki,Odds Oranı)*. https://doi.org/10.13140/RG.2.1.3786.5440

Uzgören, N., & Uzgören, E. (2015). Lisans öğrencilerinin memnuniyetini etkileyen bireysel özelliklerin istatistiksel analizi-hipotez testi, ki-kare testi ve doğrusal olasılık modeli. Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, (17).

Vijverberg, W. (2011). Testing for IIA with the Hausman-Mcfadden Test. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1882845

Williams, R. (2016). Understanding and interpreting generalized ordered logit models. *The Journal of Mathematical Sociology, 40*(1), 7-20. https://doi.org/10.1080/0022250X.2015.1112384