30th ITS World Congress, Dubai, UAE, 16-20 September 2024

Paper ID 467

Enhancing Road Safety: Insights from Delivery Drivers' Perspectives in Attica Region

Roussou Stella*[1], Konstantinos Choumis[1] , George Yannis[1]

1. Department of Transportation Planning and Engineering – National Technical University of Athens (NTUA), Athens, Greece.

*Corresponding author: s_roussou@mail.ntua.gr

Abstract

In the urban landscape of Attica, Greece, the safety of delivery drivers is a pressing issue. This study investigates their road safety challenges using a mix of qualitative and quantitative methods. Surveys and interviews with 200 food delivery drivers, alongside traffic safety data analysis, form the core of the research. Drivers responded to scenarios varying in delivery time, accident risk, and profit loss, choosing from three driving behaviour options. Using Multinomial Logistic Regression and Generalized Linear Model (GLM), key influencing factors like delivery time, accident risk perception, profit loss, driver age, fines received, and views on stricter penalties were analysed. These findings deepen our understanding of drivers' decision-making processes, aiding in devising strategies to improve road safety. This research not only offers insights and recommendations for policy and road safety enhancement but also aims to foster a safer work environment for delivery drivers, contributing to urban road safety advancements.

Keywords:
Food delivery drivers, road safety, stated preference, multinomial logistic regression, GLM analysis

1. Introduction

The landscape of urban transportation is witnessing a transformative phase globally, with the rise of e-commerce and on-demand services reshaping the way goods are delivered in cities. In Attica, a region characterized by its high population density and bustling urban activity, the role of delivery drivers has become increasingly prominent. As these drivers navigate through congested streets and complex urban terrains, their safety, as well as the safety of other road users, has emerged as a significant concern. This concern is not only a matter of individual well-being but also a public health issue, given the potential for road traffic accidents to cause serious injury or fatality. This study, focusing on the region of Attica, Greece, seeks to understand the challenges and risks these drivers face, contributing to a growing body of research in this field.

Modern society is characterized by a new era where technological development has made the delivery of food and goods even easier and more accessible for everyone than it already was [1]. In the era of mobile applications, delivery personnel work for a large number of hours and frequently use their mobile phones during deliveries for communication, GPS use, and various other reasons [1]. Despite the comfort now offered in this profession, delivery workers also face many dangers during their work. From intense and dangerous road traffic and adverse weather conditions, to the need for professional dignity and safety, delivery workers are always faced with challenges (FHWA) [2]. A significant concern is the high rate of traffic accidents involving motorcycles, which are often used by these delivery personnel. Globally, motorcycle accidents constitute a substantial percentage of fatal traffic accidents, with varying rates reported in different regions - 23% globally [3], 14% in Europe [4], 20.69% in Austria [5], and 20% in the UK [1]. In Greece, the percentage is alarmingly high at 37.66% for 2021 [6].

Our study delves into the area of urban transportation safety, with a specific focus on food delivery drivers. This area has not been extensively explored in existing research. Our research aims to fill the gap, offering a closer look at the unique challenges and risks that these drivers face in urban environments. In addition to filling a gap in the literature, this study aims to provide practical insights that can inform policy and decision-making. The unique challenges faced by food delivery drivers, such as the pressure to adhere to tight delivery schedules while navigating congested urban environments, are not only under-researched but also underrepresented in policy discussions. This research, therefore, has the potential to influence not only academic discourse but also practical strategies for enhancing road safety in urban areas.

Moreover, this study contributes to the ongoing conversation about the impact of rapid technological and service industry advancements on urban transportation systems. The proliferation of on-demand services has introduced new variables into the urban traffic equation, necessitating a re-evaluation of existing road safety strategies and policies. By focusing on a specific, rapidly evolving sector, this research provides timely and relevant insights into the broader implications of these changes for urban road safety.

This paper unfolds in a structured manner to thoroughly explore the safety challenges faced by food delivery drivers in urban settings. It begins with an introduction that outlines the study's relevance and objectives within the urban transportation safety context. The methodology section follows, detailing the questionnaire design and statistical analysis approaches. Next, the results section presents the findings from the data analysis. The paper culminates with a conclusion that summarizes the key insights, discusses their implications, and suggests directions for future research, thus offering a comprehensive perspective on the topic.

2. Methodology

This section outlines the methodology employed in this study to analyse the driving behaviour of food delivery drivers in Attica, focusing on identifying and evaluating potential risks, causes, and drivers' preferences in various hypothetical scenarios. The questionnaire development and dissemination processes are described in subsection 2.1. Based on the responses, statistical models were developed to describe driver’s choice in subsection 2.2.

2.1 Questionnaire Design and Dissemination

To gather data on the behaviour and attitudes of food delivery drivers, a structured questionnaire was developed and disseminated. The questionnaire was divided into four main parts. The first part referring to the driver background, collected data on the drivers' experience, habits, involvement in crashes, and any incidents of Road Traffic Code violations. The second part included questions to the participants about their views on the role of various factors in road safety, particularly focusing on delivery speed and its contribution to accidents. At the third and most crucial part of the questionnaire, the drivers were given hypothetical scenarios involving changes in delivery time, accident risk reduction, and profit loss per delivery. Drivers had to choose from three alternatives: a) Drive more carefully, b) Drive a bit more carefully, or c) Make no changes in their driving behaviour. At the fourth and final section of the questionnaire demographic data of the respondents were gathered, which is essential for analysing patterns and correlations in the data.

The questionnaire was distributed to 200 food delivery drivers across Attica, ensuring a mix of drivers with varying levels of experience and from different areas.

2.2 Statistical Analysis

The responses to the questionnaire were analysed using two primary statistical models:

2.2.1 Multinomial Logistic Regression

Multinomial regression analysis, a fundamental statistical technique, was employed to understand the complex relationships between several independent variables and a single dependent variable in our study. This method is particularly effective in examining the combined effect of various factors, such as driver experience, age, and number of fines, on a specific outcome, like driving behaviour or accident rates [7].

This model was employed to process the choices made by drivers in the hypothetical scenarios. This approach is suitable for scenarios where the dependent variable is categorical with more than two outcomes. In this study, it helped in understanding the factors influencing the drivers' decisions between driving more carefully, driving a bit more carefully, or making no changes in their driving behaviour.

The general form of a Multinomial regression model can be expressed as:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \cdots + \beta_\rho * X_\rho + \varepsilon \quad (1)$$

Here, Y represents the dependent variable, X_1, X_2, ..., X_p are the independent variables, β_0, β_1, β_2, ..., β_p are the coefficients of the model, and ε is the error term. The coefficients β_1, β_2, ..., β_p indicate the amount of change one can expect in Y for a one-unit change in the respective independent variable, holding all other variables constant [8].

In the context of this study, the dependent variable might be a safety-related behaviour or perception, while independent variables could include demographic factors, driving history, and attitudes towards road safety measures.

The estimation of the coefficients in Multinomial regression is typically achieved through the method of least squares [9]. This method minimizes the sum of the squared differences between the observed values and the values predicted by the model. Mathematically, the least squares criterion can be expressed as:

$$Minimize\ \Sigma\ (\ Y_i\ -\ \hat{Y}_i)^2 \quad (2)$$

Where Y_i is the observed value, $\hat{Y}_i$ is the predicted value, and n is the number of observations.
The model's effectiveness is often evaluated using the coefficient of determination, R^2, as explained by Chatterjee and Hadi (2012) [10]. This measures the proportion of variability in the dependent variable that can be explained by the independent variables in the model. The R^2 value ranges from 0 to 1, with higher values indicating a better fit of the model to the data:

$$R^2 = 1 - \frac{SS\_residual}{SS\_total} \quad (3)$$

Where SS_residual is the sum of squares of residuals and SS_total is the total sum of squares.
To ensure the reliability of the regression coefficients, standard errors are calculated, providing an estimate of the precision of the coefficients, a method discussed in detail by Kutner, Nachtsheim, and Neter (2004) [11]. A smaller standard error indicates a more precise estimate. Additionally, the statistical significance of each coefficient is typically assessed using p-values, a concept clarified by Agresti and Finlay (2009) [12]. A p-value less than a predetermined threshold (commonly 0.05) suggests that the coefficient is statistically significant.

Moreover, the application of Multinomial regression analysis in transportation research, as seen in studies like those by Hauer (1997) [13] and Washington, Karlaftis, and Mannering (2010) [14], demonstrates its utility in analysing complex, multidimensional phenomena like road safety, driver behaviour, and traffic patterns. By integrating this approach, our study aligns with contemporary research methodologies, providing robust and insightful analyses that contribute to the understanding of road safety dynamics.

2.2.2 Generalized Linear Models (GLMs)

GLM was used to analyse the relationship between the drivers' characteristics (like age, experience, number of fines, etc.) and their responses. This model is effective in handling various types of distribution for the error terms, making it versatile for analysing diverse types of data collected in the questionnaire, used in these analyses, provide a flexible and powerful framework for modelling a wide range of response variables that may not adhere to the assumptions of ordinary linear regression. GLMs extend the linear regression paradigm by allowing for different types of response variables, including binary, count, and non-normally distributed data. The key idea behind GLMs is to model the relationship between the response variable and the predictors through a systematic component called the linear predictor [15].

However, instead of assuming a linear relationship between the predictors and the response, GLMs introduce two additional components: the link function and the error distribution. The link function connects the linear predictor to the response variable, enabling the model to capture the underlying relationship. The error distribution characterizes the variability of the response and can be chosen based on the nature of the data. For example, the Poisson distribution is commonly used for count data, while the binomial distribution suits binary outcomes. GLMs allow for modelling the mean and variance of the response variable separately and can accommodate various forms of non-constant variance. Estimation of the model parameters in GLMs is typically done using maximum likelihood estimation [16, 17].

In order to analyse the results of the generalized linear model regression there is a number of factors to understand. Firstly, in regression analysis, the coefficients represent the estimated effect of each independent variable on the dependent variable. For a linear regression model, the equation is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p * X_p + \varepsilon \tag{4}$$

Where Y is the dependent variable, $X_1, X_2, \ldots, X_p$ are the independent variables, $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ are the coefficients, and ε is the error term. The coefficients measure the change in the dependent variable associated with a one-unit change in the corresponding independent variable, assuming all other variables are held constant [18]. Additionally, standard errors provide an estimate of the precision of the coefficients. They quantify the variability of the coefficient estimates across different samples. The standard error (SE) is calculated as:

$$SE(\beta_i) = sqrt[Var(\beta_i)] \tag{5}$$

Where $Var(\beta_i)$ represents the variance of the coefficient estimate. Smaller standard errors indicate greater precision and reliability of the coefficient estimates [19]. Furthermore, the significance level, or p-value, assesses the statistical significance of each coefficient. It indicates the probability of observing a coefficient as extreme as the one estimated, assuming the null hypothesis is true (i.e., no relationship between the independent variable and the dependent variable). The p-value is typically compared to a predetermined significance level (e.g., 0.05) to determine statistical significance. A p-value less than the significance level indicates statistical significance [20]. Finally, goodness-of-fit measures evaluate how well the regression model fits the data. One commonly used measure is R-squared (R^2), which represents the proportion of the variance in the de-pendent variable explained by the independent variables. R-squared ranges from 0 to 1, with higher values indicating a better fit [20].

Data Collection

The data for this study were collected through a comprehensive survey designed to capture a wide range of variables pertinent to driver behavior and road safety. The survey targeted a diverse population of drivers, encompassing various age groups, driving experiences, and geographic locations. This approach aligns with methodologies [21], ensuring a representative sample of the driving population. The questionnaire was divided into four sections, each aimed at capturing different dimensions of driver behavior, including demographics, driving habits, attitudes towards traffic laws, and accident history.

The survey was distributed using a mixed-method approach, combining both online and offline modes of distribution. Online distribution was facilitated through social media platforms, leveraging the approach recommended by Couper (2000) [22] for maximizing reach and efficiency. Offline distribution involved face-to-face interactions and paper-based surveys in various shops and offices.

Results

This research utilized two statistical models, Multinomial Regression Analysis and Generalized Linear Models (GLM), to analyse the driving behaviour of food delivery drivers. The models offer complementary insights into the factors influencing driver behaviour, encompassing aspects of safety, profit, age, frequency of fines, and attitudes towards stricter penalties.

Results from Multinomial Regression Analysis

The Multinomial Regression Analysis conducted in this study was focused on understanding the preferences of food delivery drivers across a variety of categorical outcomes, particularly their tendencies towards different driving behaviors. This analysis was pivotal in unraveling the factors that influence the decisions and actions of these drivers on the road. The data for this analysis derive from the choice between three options for the hypothetical scenarios which were given to the respondents. The key findings from the multinomial regression analysis are demonstrated in Table 1, where choice 1 stands for the case of driving more carefully and choice 2 drive a little bit more carefully. Both Choices are compared to the case of Choice 0, which represents the scenario of no change.

Table 1: Results from the Multinomial Regression Analysis

Choice=1	**coef**	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
const	-40.834	14.582	-2.800	0.005	-69.413	-12.255
Time_norm	23.775	335.512	0.071	0.094	-633.817	681.367
AccidRed_norm	57.262	118.214	0.484	0.063	-174.434	288.957
Profit_norm	28.835	284.166	0.101	0.092	-528.121	585.790
AGE	0.358	5.180	0.069	0.945	-9.795	10.512
TIMES_FINE	-0.165	2.162	-0.076	0.939	-4.402	4.073
STRICT_PENALTIES	0.444	6.676	0.067	0.947	-12.640	13.529

Choice=2	**coef**	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
const	-7.2989	10.5800	-0.690	0.490	-28.0350	13.4370
Time_norm	13.0971	335.5110	0.039	0.0970	-644.4920	670.6860
AccidRed_norm	12.9466	117.0410	0.111	0.0910	-216.4500	242.3440
Profit_norm	24.8793	284.1650	0.088	0.0930	-532.0740	581.8330
AGE	-0.4412	5.1760	-0.085	0.9320	-10.5850	9.7030
TIMES_FINE	-0.0417	2.1590	-0.019	0.985	-4.2740	4.1910
STRICT_PENALTIES	-0.0059	6.6710	-0.001	0.999	-13.0800	13.0680

A significant result from the analysis was the strong preference among drivers for cautious driving, closely linked to an emphasis on increased safety measures. This was evident from the high coefficients associated with safety variables within the model. Such a finding indicates that reducing the risks of accidents is a major motivating factor for drivers to adopt careful driving practices. It reflects an inherent understanding and concern among drivers regarding the importance of safety, not just for themselves but also for the broader community.

Interestingly, the results demonstrated a tendency among drivers to prioritize safety over profit. This suggests that drivers are aware of the long-term benefits of safe driving practices, which could overshadow potential short-term financial losses. The inclination to choose safety over immediate economic gain highlights a responsible attitude among drivers, recognizing that the benefits of cautious driving extend beyond monetary considerations.

The analysis also shed light on the influence of age on driving behavior. Older drivers showed a higher propensity for cautious driving, which could be attributed to their accumulated driving experience and potentially more mature risk assessment capabilities. This finding suggests that with age and experience, drivers tend to become more aware of the risks associated with driving and are more inclined to adopt safer driving habits.

An intriguing aspect of the findings was the behavior of drivers with a history of receiving more fines. This group of drivers tends to engage in riskier driving behaviors, which could indicate a desensitization to the penalties or a habitual non-compliance with traffic laws. This observation points towards the possibility that for some drivers, fines and penalties might not serve as effective deterrents against risky driving behaviors, suggesting a need for more targeted interventions for this particular group.

Results from Generalized Linear Models (GLM)

The application of Generalized Linear Models (GLM) in our study provided critical insights into the likelihood of drivers choosing safer driving practices under various conditions. The GLM approach was particularly useful in quantifying how certain factors influenced the probability of making specific driving choices. The data used for the implementation of GLM models were derived from three questions. The first one is about the strictness of measures in order to reduce traffic accidents (Table 2) , the second one about the fines increase in order to improve road safety (Table 3) and the last one about the camera uses to monitor violations in order to improve road safety (Table 4).

Table 2: GLM Model for strict measures

	coef	std err	z	P>|z|	[0.025	0.975]	VIF
const	-0.184	0.189	-0.973	0.330	-0.553	0.186	13.628
HELMET_USE	0.724	0.144	5.035	0.000	-0.442	1.006	1.070
MOTO_EXP	-0.084	0.042	-2.014	0.044	-0.166	-0.002	1.054
TIMES_SERIOUS_ACCID	0.163	0.082	1.986	0.047	0.002	0.323	1.111
PROH_SIGN_VIOLATION	-0.337	0.162	-2.081	0.037	-0.655	-0.020	1.070
TIMES_FINE	-0.056	0.042	-1.336	0.181	-0.137	0.026	1.104

Table 3: GLM Model for increase fines

	coef	std err	z	P>|z|	[0.025	0.975]	VIF
const	0.6572	0.1550	4.2430	0.0000	0.3540	0.9610	8.7878
SUIT_USE	1.0403	0.1200	8.6590	0.0000	0.8050	1.2760	1.0285
MOTO_EXP	-0.1665	0.0440	-3.7950	0.0000	-0.2530	-0.0810	1.1099
TIMES_FINE	0.1042	0.0430	2.4350	0.0150	0.0200	0.1880	1.1008
TIMES_ACCID	-0.2160	0.0550	-3.9620	0.0000	-0.3230	-0.1090	1.1273
RED_LIGHT_VIOLATION	-0.0199	0.2230	-0.0890	0.9290	-0.4570	0.4170	1.0296

Table 4: GLM Model for cameras use

	coef	std err	z	P>|z|	[0.025	0.975]	VIF
const	1.1185	0.2720	4.1170	0.0000	0.5860	1.6510	19.1426
MOTO_EXP	-0.0901	0.0550	-1.6520	0.0990	-0.1970	0.0170	1.1652

HELMET_USE	0.3487	0.1640	2.1210	0.0340	0.0270	0.6710	1.0303
TIMES_FINE	-0.0417	0.0490	-0.8440	0.3980	-0.1390	0.0550	1.0446
WORK_TIME	0.1784	0.0930	1.9200	0.0550	-0.0040	0.3610	1.1593
PASS_BY_VIOLATION	-0.5673	0.1680	-3.3770	0.0010	-0.8970	-0.2380	1.0431

The GLM results revealed a significant relationship between drivers' attitudes toward road safety and their driving behaviors. Specifically, drivers who were in favor of stricter penalties exhibited a higher likelihood of engaging in cautious driving behaviors. This finding indicates an acknowledgment among these drivers of the necessity for more robust road safety measures. The positive correlation between the support for stricter penalties and safer driving practices suggests that attitudes towards road safety regulations play a crucial role in influencing driver behaviour.

The GLM analysis corroborated the findings from the Multinomial Regression Analysis, showing a clear age-related trend in driving behaviors. Older drivers demonstrated a greater propensity towards safer driving practices. This trend is an important insight, as it highlights the influence of age and possibly accumulated driving experience on risk assessment and decision-making on the road. The GLM provided additional clarity on the extent of this influence, quantifying the increasing likelihood of safer driving practices with advancing age.

The GLM results also shed light on the complex interplay between economic motivations and safety considerations. The analysis suggested that economic incentives or disincentives have significant impacts on driving behaviors. Drivers seemed to weigh profit motives against safety concerns, indicating that financial considerations are a crucial factor in their decision-making process. This finding underscores the need to balance economic incentives with safety to encourage more responsible driving behaviors.

Another nuanced insight from the GLM analysis pertains to the influence of fines on driving behaviour. The results highlighted that drivers with a history of frequent fines are less likely to modify their behaviour towards safer practices. This trend might indicate a habitual disregard for traffic regulations among these drivers. The correlation between the frequency and severity of fines and the likelihood of adopting safer driving habits provides an important perspective on the effectiveness of penalties in modifying driver behaviour.

In conclusion, the GLM analysis offered a comprehensive understanding of the factors influencing food delivery drivers' choices towards safer driving practices. By analysing attitudinal factors, age-related trends, economic considerations, and the impact of fines, the study provides valuable insights for developing targeted strategies to promote safer driving behaviors among this group.

Conclusions

The combined results from both Multinomial Regression Analysis and GLM provide a comprehensive overview of the factors influencing food delivery drivers' behaviour. Safety concerns emerge as a primary factor, with notable variations based on age and personal experience with traffic incidents. The relationship between the frequency of fines and driving behaviour underscores the need for more effective enforcement and awareness campaigns. Additionally, the positive response to stricter penalties points towards a potential avenue for policy intervention to enhance road safety. These findings offer valuable insights for stakeholders aiming to improve traffic safety and influence driver behaviour in urban settings.

This study, through Multinomial Regression Analysis and Generalized Linear Models (GLM), has provided critical insights into the driving behaviors of food delivery drivers, bearing significant implications for policy-making, driver education, and public safety initiatives. A standout observation from the study is the drivers' tendency to prioritize safety over profit. This finding is crucial for policymakers and the food delivery industry, as it underscores the need to create a working environment where drivers are not forced to choose between their safety and earnings. Implementing policies that reward safe driving practices could further reinforce this positive trend.

Age and experience play a significant role in driving behaviour, with the analysis indicating a propensity for older drivers to engage in safer driving practices. This insight suggests the effectiveness of targeted training programs for younger or less experienced drivers, focusing on risk assessment and the promotion of safe driving techniques. The study also reveals an intriguing aspect of how penalties influence driving behaviour. The observed trend where drivers with a higher number of fines tend to engage in riskier behaviour points to a potential desensitization to penalties. This finding suggests a need for a re-evaluation of the penalty system, with a shift towards more educational and rehabilitative measures, rather than relying solely on punitive actions.

Furthermore, the study highlights the complex interplay between economic pressures and driver safety. It is crucial for companies and policymakers to consider how economic factors impact driving behaviors and to design compensation and incentive structures that encourage safe driving practices. Additionally, the study finds a correlation between attitudes towards stricter penalties and safer driving behaviors. This suggests that initiatives aimed at transforming driver attitudes, perhaps through awareness campaigns and community involvement, could play a significant role in promoting safer driving practices.

One of the primary limitations is the sample diversity. The study predominantly focuses on male delivery workers aged between 25 and 50 years. This concentration limits the generalizability of the findings to other demographic groups, such as female delivery workers or those outside this age range. Additionally, the findings are based on a specific geographical area. As such, the results may not fully represent the experiences of delivery workers in different regions or countries with varying traffic conditions and cultural attitudes towards road safety. Furthermore, the reliance on self-reported data in the survey can introduce biases, as participants may underreport behaviors like traffic rule violations or over-report incidents like accidents due to social desirability or memory recall issues.

Future research should aim to include a more diverse group of participants. Incorporating a wider age range, more female workers, and different socio-economic backgrounds would provide a more holistic understanding of the delivery workforce. Long-term studies could offer insights into how the job conditions and behaviors of delivery workers evolve over time, especially in response to changes in the gig economy and traffic regulations. Finally, research into the effects of new technologies (e.g., navigation aids, safety apps) and policy interventions (e.g., stricter traffic laws, improved worker rights) on the safety and well-being of delivery workers would be valuable.

References

1. Christie, N., & Ward, H. (2023). Delivering hot food on motorcycles: A mixed method study of the impact of business model on rider behaviour and safety. Safety Science, 158. https://doi.org/10.1016/j.ssci.2022.105991.
2. FHWA. Travel time reliability: Making it there on time, all the time.
3. WHO, New global guidelines to curb motorcycle crash deaths, 2022.
4. European Commission-European Commission. (2020). 2019 Road Safety Statistics: What Is behind the Figures?
5. Eichberger, A., Kraut, M., & Koglbauer, I. v. (2022). Improved Perception of Motorcycles by Simulator-Based Driving Education. Sustainability (Switzerland), 14(9). https://doi.org/10.3390/su14095283
6. NRSO, National Technical University of Athens, Road Safety Observatory
7. El-Habil, A. M. (2012). An application on multinomial logistic regression model. Pakistan journal of statistics and operation research, 271-291.
8. Wooldridge, J.M. (2009) Introductory Econometrics: A Modern Approach. 4th Edition, South-Western College Publishing, Cengage Learning, Boston, 22-68.
9. Draper, H.R. and Smith, H. (1998) Applied Regression Analysis. 3rd Edition, John Wiley & Sons Inc., New York, 713. http://dx.doi.org/10.1002/9781118625590.
10. Chatterjee, S. and Hadi, A.S. (2012) Regression Analysis by Example. 5th Edition, Wiley, New York, 98.
11. Applied Linear Regression Models- 4th Edition with Student CD (McGraw Hill/Irwin Series: Operations and Decision Sciences) 4th Edition by Michael Kutner (Author), Christopher Nachtsheim (Author), John Neter.
12. Statistical Methods for the Social Sciences (4th Edition) 4th Edition by Alan Agresti, Barbara Finlay.
13. Hauer, E. (1997) Observational Before-After Studies in Road Safety—Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety. Elsevier Science Incorporated, Tarrytown, NY.

14. Statistical and Econometric Methods for Transportation Data Analysis, By Simon P. Washington, Matthew G. Karlaftis, Fred Mannering.
15. McCullagh, P., & Nelder, J. A. (1989). Generalized linear models CRC press. In Boca Raton. USA: Florida.
16. Cameron, A. C., & Trivedi, P. K. (2013). Regression analysis of count data (Vol. 53). Cambridge university press.
17. Hardin, J. W., & Hilbe, J. M. (2007). Generalized linear models and extensions. Stata press.
18. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introduction to linear regression analysis. John Wiley & Sons.
19. Greene, W. H. (2018). Econometric Analysis (N. Bhalla, Ed.)
20. Wooldridge, J. M. (2015). Introductory econometrics: A modern approach. Cengage learning.
21. How to conduct your own survey, Priscilla Salant and Don A. Dillman, 1994, https://api.semanticscholar.org/CorpusID:62697981
22. Couper, M. (2000) Web surveys: A review of issues and approaches. Public Opinion Quarterly, 64, 464-494. http://dx.doi.org/10.1086/318641