International Journal on Advanced Science Engineering Information Technology

Vol.15 (2025) No. 1
ISSN: 2088-5334

Shipment Size Analysis on Cross-Border Freight Transportation Using Stereotype Logistic Regression: Case Study Between West Kalimantan, Indonesia and Sarawak, Malaysia

Said Basalim [a,*], Danang Parikesit [b], Sigit Priyanto [b], M. Zudhy Irawan [b]

[a] *Department of Civil Engineering, Faculty of Engineering, Tanjungpura University, Pontianak, Indonesia*
[b] *Department of Civil and Environmental, Faculty of Engineering, Gadjah Mada University, Yogyakarta, Indonesia*
Corresponding author: [*]saidbasalim@gmail.com

***Abstract*—Cross-border freight transport is vital for regional logistics, facilitating trade between neighboring countries. However, inefficiencies in shipment size decisions, such as overloading or underutilizing vehicles, can lead to increased costs and logistical challenges. This study investigates the key factors influencing shipment size decisions in the West Kalimantan (Indonesia)–Sarawak (Malaysia) cross-border corridor. Data from 2,017 truck trips conducted between 2016 and 2018, provided by the Fish Quarantine and Inspection Agency, was used for the analysis. Ordinal logistic regression methods were employed to examine the relationship between shipment size and influencing variables, including company type, commodity value, transport distance, and vehicle type. Shipment sizes were categorized into five groups, ranging from less than 500 kg to more than 7,500 kg. The results indicate that the stereotype logistic regression model better fits the data than the generalized ordered logit model based on log-likelihood, R-squared, and information criteria values. Key findings highlight that firm classification and commodity value significantly impact shipment size decisions, while transport distance and vehicle type have a less pronounced effect. This study emphasizes optimizing shipment size decisions to enhance cross-border transport efficiency and reduce costs. The findings offer valuable insights for policymakers and logistics companies to develop targeted strategies, especially in the context of marine commodity exports. The study improves regional trade competitiveness and promotes sustainable logistics practices by addressing inefficiencies in shipment size decisions.**

***Keywords*—Commodity value; cross-border; freight transport; ordinal logistic regression; shipment size; stereotype logistic regression.**

Manuscript received 4 Aug. 2023; revised 17 Dec. 2024; accepted 18 Jan. 2025. Date of publication 28 Feb. 2025.

I. Introduction

Cross-border freight transportation is a key aspect of transportation science and plays a vital role in regional trade and facilitating bilateral exchanges between neighboring countries with shared land borders. This activity involves various operations and processes that enable the delivery of goods across national boundaries [1]. However, inefficiencies in managing shipment sizes, such as overloading or underutilization, impose substantial economic burdens. The West Kalimantan (Indonesia)–Sarawak (Malaysia) corridor is a key trade and freight transportation route, causing a detailed analysis of how shipment size decisions affect the efficiency of freight operations. Trucks and trailers are the most used modes of transport in regions such as Europe and Asia, where direct land borders exist. For instance, the Europe-Turkey border is a critical trade corridor [2], [3]. Research has found key cost components in freight transportation, including transaction, policy, time, and transportation costs [4]. Past studies, for example, emphasize that cross-border freight volume doubled between 1995 and 2015, raising concerns about environmental impacts and the need for mitigation policies [5]. Likewise, challenges in e-commerce planning, such as prolonged transport times and low service levels, emphasize the importance of optimization in warehouse operations [6].

Although inventory stages typically have a minimal impact on overall logistics costs, in the context of cross-border exports—particularly for commodities such as fresh seafood that dominate the research site—inventory management can lead to substantial cost surges [7], [8]. This system is particularly relevant for specific commodities, such as fresh fish, which dominate cross-border exports in the study area. Fresh seafood is transported directly from fishing ports or cultivation sites to receivers. The initial step in optimizing shipping efficiency is to examine the shipment, particularly its size.

Several studies discussing less-than-truckload shipments include [9], [10]. If the economies of scale prompt the company to adopt the policy to carry larger quantities of freight, several considerations, such as distance and flexibility, can lead to different decisions being made. Freight transportation decision analysis dates back several decades, starting with Nam and Ki-Chan in the late 1990s and continuing with Combes, Samimi, Mohammadian, and Kawamura in the early 2000s [11].

The existing body of research in freight transportation primarily focuses on the choice of transportation modes within the shipping model. Shipment mode choice analysis is typically conducted using the microeconomic discrete choice method, involving a set of transportation alternatives and predefined goals [12], [13]. This alternative is characterized by the utility function representing the value of each option from the shippers' perspective. Shipment size is analyzed as a decision unit in modeling, bridging the gap between traffic modeling methods and companies' logistics behavior.

The notion of shipment size refers to the quantity of items transported simultaneously from the sender or shipper to the recipient [13], [15]. Despite its importance in logistics, research on this topic remains limited. In this context, the term "lot size" is often used, indicating that each shipment delivered to the receiver or distributed between depots and distributors has a consistent quantity [16]. The total number of shipments for specific commodities can be determined by multiplying the lot size by shipment frequency. To address these challenges, [17] proposed a simplified goods distribution model to estimate the weighted average delivery distances at various supply chain phases. According to [18] the movement of goods typically involves four stages: assessing the volume of products being sent and received, identifying proper suppliers and recipients, and determining the mode of transportation. However, route planning during these processes often overlooks shipment size, which can affect efficiency. Research findings by [19], indicate that shipment size is influenced by the nature of the commodity (e.g., agricultural or wood products) and the specific characteristics of its users (e.g., offices, restaurants, or factories). Furthermore, [20] highlighted that optimizing shipment size involves balancing order quantity, reorder levels, and production rates to minimize total costs and emissions under conditions of random demand, backorders, and potential sales losses. A study by [21], [22] investigated how route changes, transportation distances, carrier company characteristics, and transportation modes influenced the selection of the optimal vehicle size. In the present study, the authors [13] developed a model to analyze the relationship between delivery and vehicle size. The findings indicated that the primary factors influencing vehicle size choice are the expenses associated with vehicle operation, the vehicle's age, and the operator's characteristics. Additionally, factors such as the distance covered, overall shipment demand, item nature, and the cargo density significantly affect shipment size.

Initially treated as continuous data, shipment size was subsequently categorized during the analytical phase to facilitate a more nuanced understanding. This categorization was based on numerous factors, including the observed minimum and maximum shipment weights. Shipment size was categorized into five categories as described [23]: 3,500 kg, 3,501-15,000 kg, 15,001-30,000 kg, 30,001-100,000 kg, and >100,000 kg. Similarly, [24] segmented shipment sizes into four categories: 1-10 kg, 10-100 kg, 100-1,000 kg, and above 1,000 kg.

Analyzing and modeling shipment size decisions is challenging due to limited data and a lack of understanding of the microeconomic principles underlying company decisions. A comprehensive analysis of cross-border freight transport indicates that only one mode of transport is used for exports: trucks. Tracks are the only viable option, as no other modes of transport can compete with their cost-effectiveness, flexibility, and ability to handle the large volumes of cargo required for this trade.

A comprehensive literature review has identified key factors influencing cargo size variance. According to [14], these factors include transport distance, demand for shipping services, commodity type, and cargo density. [15] asserts that the patterns of export shipments are influenced by multiple factors: the export potential of a product in a specific market, the revenue generated per shipment, and the firm's efficiency in managing these shipments. These elements collectively influence strategic decisions regarding shipment sizes.

In reviewing the number of shipped commodities or shipment sizes, two conditions using trucks—full truckload (FTL) and less than truckload (LTL)—are considered. FTL is a shipment mode for suppliers to rent and use the entire truck to ship packages, whereas LTL is a shipment mode for suppliers that outsource shipment to a third party based on demand [16]. This study examines the factors influencing shipment size decisions in cross-border freight transport, with a focus on marine commodities. It identifies key independent variables that impact shipment size, using ordinal logistic regression and its variants for analysis. Efficient export of commodities, particularly fresh seafood, is vital for increasing local revenue and optimizing logistics costs. A major issue is the inefficiency of using large trucks to transport low-tonnage goods, caused by a mismatch between shipment size and transport capacity, which can result in higher transportation costs. Despite the importance of this issue, limited research has addressed shipment size decisions in cross-border freight transport, particularly in Indonesia. Existing studies have largely focused on mode choice and alternative transportation methods, often neglecting the complexities of shipment size. This research aims to fill this gap, offering insights into optimizing cross-border freight transport.

The primary aim of this study is to examine the factors influencing shipment size in cross-border freight transportation, particularly focusing on marine commodities, and to identify the decision-making preferences of companies regarding shipment size. This research investigates the trade-offs between logistics costs and inventory costs, assessing how larger shipments, which incur higher inventory costs, compare with smaller, more frequent shipments that result in higher logistics costs. The analysis is essential not only for addressing inefficiencies when large-capacity trucks are used for low-tonnage shipments but also for enhancing the efficiency of cross-border logistics, thereby minimizing unnecessary costs that directly impact export revenues. Furthermore, this study provides insights that advance the understanding of logistical efficiencies and serves as a crucial basis for policymaking to enhance economic cooperation and sustainable trade practices between Indonesia and Malaysia.

II. MATERIALS AND METHODS

A. Construction of the variables from choice-based data

This study investigates the export dynamics of marine products along the cross-border freight corridor connecting West Kalimantan, Indonesia, and Sarawak, Malaysia, with a focus on the unique characteristics of the cross-border freight transportation system. The dataset comprises detailed records from the Entikong Fish Quarantine and Inspection Agency (BKIPM), covering the period from 2016 to 2018. This dataset was selected for its comprehensive coverage of cross-border shipments and underwent rigorous cleaning and preprocessing before analysis to ensure accuracy and relevance to the study's objectives. The BKIPM is located 5.4 km from the border crossing point. This corridor accounts for 99% of marine commodity exports to Malaysia (Central Statistics Agency of West Kalimantan, 2018).

Shipments with vehicles bearing Indonesian license plates are directed to the Inland Port of Tebedu in Sarawak, Malaysia, as trucks from West Kalimantan are prohibited from proceeding further into Malaysia. At this stage, the marine commodity cargo, destined for buyers or customers, is transferred to Malaysian trucks for the remainder of the journey. The analysis classifies the commodity groups according to the Harmonized System Code (HS Code). Using the disaggregate approach, daily information related to shipments, commodities, and transportation modes [17] includes the origin trip location and haul length.

Since the transported commodity is fresh marine produce and the preservation method involves packing the fish in boxes with ice (a conventional method), all trucks, after receiving the commodity from the nearest fish port, travel directly from the export loading point to the border, and then to the Tebedu Inland Port (Fig. 1). Haul length refers to the distance traveled by trucks from the exporter's office (fish processing unit) to the destination.

Fig. 1 Cross-border corridor freight transportation between West Kalimantan, Indonesia, and Sarawak, Malaysia

The primary data used in this study comprise daily records of truck movements across borders, transporting marine commodities over three years from 2016 to 2018. Data from 2019 to 2021 were excluded due to disruptions caused by the COVID-19 pandemic, which halted most exports by crossing the border. This study addresses the potential biases introduced by this exclusion and employs statistical methods to mitigate their impact on the research findings. The dataset comprises records from 2,017 export trips facilitated by eleven companies, utilizing four vehicle types to transport five categories of commodities, classified according to the Harmonized System (HS) Code. Cross-border freight transportation at the West Kalimantan and Sarawak border is primarily served by trucks. Some commodities shipments have been recorded using pickups. Marine commodities are transported using three freight vehicles: small freight vehicles, medium-sized box trucks, and large trucks.

The daily data collected include several key variables: the names of shipping companies, HS codes, shipment weights, commodity values, and types of freight vehicles used (Table I).

TABLE I
VARIABLES USED IN SHIPMENT SIZE ANALYSIS

Variable		Definition
Company type	*ltdp*	0: sole proprietorship 1: limited partnership
Type of freight vehicle	*tpfv*	1: small freight vehicles 2: medium box-trucks 3: large trucks
Value of commodity	*comv*	In 1,000 IDR
Weight of commodity	*comw*	In 1,000 kg
Classification of commodity	*comc*	1: shrimp, crab 2: fish 3: combination
Distance (length of haul) (km)	*lgth*	Continuous data; suggested distance between origin and destination by Google maps (km)
Distance (length of haul) (km)	*lgth*	Continuous data; suggested distance between origin and destination by Google maps (km)
Response		
Shipment size in kg	*shps*	1: ≤ 500 (very low) 2: 501 – 2,500 (low) 3: 2,501 – 5,000 (medium) 4: 5,001 – 7,500 (high) 5: > 7,500 (very high)
Number of shipments	*nosh*	

Data from the BKIPM are crucial for assessing Indonesia's cross-border transport system. They offer valuable insights into cargo sizes, typically challenging to obtain. The BKIPM provides comprehensive daily data on truck-exported marine commodities, including the arrival date and time, exporter names and types, commodity names, classifications (HS Codes), weights, values, and types of freight vehicles, along with their license plates. Building on prior studies, shipment sizes in this research were stratified into five categories based on the weight range of commodities, which ranges from less than 500 kilograms to more than 7,500 kilograms.

B. Methods

The research question is derived from the Background section, which outlines the factors export commodity shippers consider when determining shipment size. This question is addressed by examining the relationship between the independent variables (influencing factors) and the dependent variable (shipment size). Quantitative research methods, including correlational and causal-comparative techniques [18], were applied to analyze the survey data.

Various survey methods are available for collecting data on truck trips, including those transporting maritime commodities for export to neighboring countries, as utilized by [19], [20]. This study collected data from cross-sectional preference surveys conducted over three years [19]. Interviews were conducted with truck drivers at 27 internal checkpoints and seven external checkpoints around the city to collect data. Similarly, [21] collected information from a business travel survey, which included shipment details such as item type and destination, as well as company characteristics, such as industry type and number of employees.

For this study, the researcher was fortunate to obtain data from the Fish Quarantine and Inspection Agency (BKIPM) near the international border. The acquired data pertains to company characteristics, commodity types, and commodity weights. Obtaining this data was considerably more efficient than conducting roadside or city boundary surveys at various locations in West Kalimantan Province. In addition to the data above, information on vehicle types was collected through direct observations of vehicles undergoing administrative procedures at the Quarantine Agency and border posts. The study population comprises all recorded maritime commodity export activities documented by BKIPM. This focus was selected due to the significant economic activities and logistical dynamics associated with this specific trade route. The daily freight vehicle travel data collected over three consecutive years represent an invaluable sample. This sample provides a comprehensive and continuous dataset. The three-year span allows for in-depth trend analysis and a better understanding of cross-border transportation dynamics. Success in collecting daily data ensures accuracy in the predictive models and statistical analysis employed, underscoring the reliability of the findings of this study. With extensive data coverage, this sample can be considered representative of freight transportation behaviors in the studied region, enhancing the relevance and broader applicability of the study's findings.

To analyze the relationship between independent and dependent variables, particularly for shipment sizes with an ordered structure, we employed an ordinal logistic regression model appropriate for handling ordered categorical data. Before applying the model, we verified key assumptions, including the absence of multicollinearity among predictors and the proportional odds assumption across response categories. As described previously, shipping sizes are categorized into five ordered groups, ranging from the heaviest to the lightest. This model estimates the probability of a shipping size falling into each ordered category based on the influence of the independent variables.

Using a disaggregated methodology, we examined transportation choices by individual entities responsible for logistical decisions. This approach provides detailed insights into the factors influencing the selection of transportation modes, routes, and shipment volumes, offering a more profound understanding than the aggregate approach. The disaggregate approach provides deeper insights into individual decision-making processes than the aggregate approach, enhancing our knowledge of logistical preferences and strategies.

To analyze freight mode choice, we used an econometric model that expresses this choice as a function of key independent variables, providing a nuanced and detailed understanding of transportation dynamics. By applying a discrete-continuous goods option model to deliver data across various commodity types, we integrated independent variables and weighting schemes to conduct a comprehensive analysis of freight transportation choices [22].

Logistic regression, a widely used statistical technique, is employed to analyze the relationship between response and predictor variables [23]. In shipment size analysis, several alternative models may be considered. Guidelines for model selection in cross-border shipment size export analysis can be developed based on the information provided in [24], as outlined below:

a. The ordinal least squares regression method will be used if the dependent variable is continuous or categorical with more than five levels, which can be treated as continuous variables.
b. If the dependent variables are ordered, then the ordered logistic regression method will be applied.
c. The tests are necessary if the dependent variables are suspected to be unordered. Two tests, namely the *omodel* and Brant test, are conducted. If the proportional odds assumption of the ordered logistic regression is violated, two discrete choice models could be used: generalized ordered logit and stereotype logistic (SL) regression.

Fitstat is a statistical tool in STATA that calculates various fit measures for regression models during the post-estimation phase [25]. These measures include log-likelihood, chi-square, R^2, and information criteria (IC). Ordinal or ordered logistic regression can be analyzed using various statistical software, including proprietary software such as SPSS, SAS, and STATA, as well as open-source software like R [26], [27], [28]. The development of STATA was underpinned by the need for software capable of analyzing generalized ordered logistic regression and stereotype logistic regression.

The three main types of discrete choice models are the logit, probit, and multinomial logit models [29]. The logit model remains the most widely used and straightforward due to its easily interpretable choice probability formula. Logistic regression is a statistical technique used to model the relationship between a dependent variable (the outcome or response variable) and a set of independent variables (predictors or explanatory variables) [30].

C. Ordinal Logistic Regression

In some cases, the scales derived from the outcome categories are ordinal rather than nominal [31]. These cases require shifting from a multinomial approach to an ordered logistic regression methodology. For example, in this study, shipment size is classified into five categories: very low, low, medium, high, and very high.

Ordinal regression models are closely related to logistic regression models for dichotomous (binary) outcomes in specific statistical software analyses [32]. In research employing ordered logistic regression, several types of comparisons can be performed, including cumulative, sequential, and adjacent comparisons, with the cumulative type corresponding to parallel odds (PO) [33], [34]. When the dependent variable is an ordinal scale representing a continuous underlying measure, such as income intervals, the most appropriate approach is to use the cumulative method.

For modeling shipment size, this study utilizes continuous data categorized into five response levels, as shown in Table I. According to [33], the ordinal logistic regression model assumes that the outcome variable is a latent variable, expressed in logit form as shown in Eq. (1).

$$\text{logit}[\pi(Y_i > j \mid x_1, x_2, .., x_p) = ln\left(\frac{\pi(Y \le j \mid x_1, x_2, ..., x_p)}{\pi(Y > j \mid x_1, x_2, .. x_p)}\right) = \alpha_j + (-\beta_1 X_1 - \beta_2 X_2 - \cdots - \beta_p X_p) \quad (1)$$

where $\pi j(x) = \pi\ (Y \le j \mid x1, x2, \ldots, xp)$, which is the probability of being at or below category j, given a set of predictors j = 1,2, …, J – 1. *Aj* denotes the cut points, and β1, β2,…, βp are the logit coefficients.

D. Generalized Ordered Logit

The Generalized Ordered Logit model was developed by [35] to address specific cases, namely the Proportional Odds (PO) model (also known as the parallel lines model) and the partial PO model. With this capability, the Generalized Ordered Logit provides an alternative solution for ordinal dependent variables that do not satisfy the PO assumption. The gologit2 equation, proposed by Fu in 1998 [35], is as follows:

$$logit[\pi(Yi > j \mid = x_1, x_2, . . xp)| = \ln\left(\frac{\pi(Y > j \mid x1, x2, \ldots, xp)}{\pi(Y \le j \mid x1, x2, \ldots, xp)}\right) = \alpha j + (\beta 1jX1 + \beta 2jX2 + \cdots + \beta pjXp) \quad (2)$$

E. Stereotype Ordered Logit (SL)

J.A. Anderson introduced the Stereotype Ordered Logit (SL) model in his paper, stating that this model can identify constant and unchanging outcomes by categorizing the intensity levels of the subject under study. The SL model is an alternative approach that treats the response variable as categorical rather than ordinal, especially when there is uncertainty regarding the relevance of the ordering in the response variable [37].

The relationship between the SL model, multinomial logit, and proportional odds (PO) is explored in [38]. Additionally, [33] proposed expressing Anderson's stereotype ordinal regression model in the form given in Eq. (3), as follows:

$$logit\ [\pi(j, J)] = \ln\left(\frac{\pi(Y = j \mid x_1, x_2, ..., x_p)}{\pi(Y = J \mid x_1, x_2, ..., x_p)}\right) = \alpha_j - \phi_j(\beta_1 X_1 + \beta_2 X_2 + \ldots \beta_p X_p) \quad (3)$$

where j = 1,2, …, J – 1; J is the baseline or reference category, the last category here, but can be any other category chosen by the researcher. Y is the ordinal response variable with categories from j to J; αj represents the intercepts. $\beta_1, \beta_2, \ldots, \beta_p$ are logit coefficients for the predictors, X_1, X_2, …, X_p, respectively. The ϕ_j represents the constraints used to confirm whether the outcome variable is ordinal, provided that the condition expressed in Eq. (4) is satisfied [38], as follows:

$$1 = \phi_1 > \phi_2 > \phi_3 > \cdots \phi_{J-1} > \phi_J = 0 \quad (4)$$

III. Results and Discussion

A. Logistic Regression Analysis

At this stage, the *ologit* command in a STATA program is used to obtain the outcomes of an ordered logistic regression model. In Table II, seven variables are displayed in two separate sections to ensure optimal readability of the data.

TABLE II
LIST 7 DATA FROM 2,017 ROW DATA MARINE COMMODITIES CROSS-BORDER EXPORT BETWEEN WEST KALIMANTAN AND SARAWAK

Number of shipments	Company type (*ltdp*)	Distance (*lgth*; km)	Value of comm. (*comv*; millions of IDR)	Classification of comm. (*comc*)	Weight of comm. (10 kg)	Type of freight vehicle	Shipment size (kg)
1	1	399	93.75	3	270.0	2	3
2	0	248	85.2	1	210.0	2	2
3	0	246	85.05	3	217.5	2	2
4	0	248	57.75	3	165.0	2	2
5	0	248	79.05	3	213.0	2	2
6	1	399	82.38	3	237.0	3	2
7	1	399	93.75	3	270.0	2	3

The notation follows the format presented in Table I in the previous section. The modeling results obtained through ordered logistic regression are subsequently tested for parallel odds (PO) assumptions using the *omodel* command in STATA. The Brant test cannot be conducted on the research data because not all independent variables can be included in the binary logits. Therefore, the *omodel* test employs the likelihood ratio test. The null hypothesis posits no difference in the coefficients between the models. Consequently, the test results should be statistically insignificant, as indicated by a Prob > $\chi 2$, with a value less than 0.001. The outcomes of the *omodel* test are presented in Table III.

TABLE III
THE RESULT OMODEL LOGIT

Approximate likelihood-ratio test of odds across response categories	
χ^2 (15)	287.01
Prob > χ^2	0.0000

The violation of the proportional odds assumption (for ordered logistic regression) in the model is confirmed by the significant of Prob> $\chi2$ value of 0.000. Furthermore, *gologit2* and *slogit* are alternative methods for modeling cross-border shipment scenarios. Comparison and analysis of the *gologit2* and *slogit* modeling outcomes are performed using the *fitstat* command in STATA.

TABLE IV
SELECTION OF THE APPROPRIATE MODEL WITH FITSTAT (GOLOGIT2 AND SLOGIT MODELS)

	gologit2	**slogit**
Log-likelihood		
Model	-1385.570	-1308.166
Intercept-only	-2153.324	-2153.324
Chi-square		
Deviance (df=2008)*; (df=2004)**	2771.140*	2616.333
p-value	0.000	0.000
R2		
McFadden	0.357	0.392
McFadden (adjusted)	0.352	0.386
Cox-Snell/ML	0.533	0.567
Cragg-Uhler/Nagelkerke	0.604	0.644
IC		
AIC	2789.140	2644.333
AIC divided by N	1.383	1.311
BIC (df=9)*; (df=14)**	2389.625*	2722.864**

Based on the results shown in Table IV, and through a comparison of the *parameters*: (i) log-likelihood, (ii) $\chi2$, (iii) R2, and (iv) IC, the most suitable model for shipment size in cross-border freight transportation between West Kalimantan, Indonesia, and Sarawak, Malaysia, is identified as the SL model (*slogit*). Table V presents the outcomes of the SL regression model using STATA.

TABLE V
RESULT OBTAINED USING THE STEREOTYPE LOGISTIC REGRESSION MODEL

The number of obs.					**2,107**
Model		**Wald χ^2 (7)**		126.66	
		Prob > χ^2		0.0000	
		Log-likelihood		-1308.1663	
Coeff.					
ltdp	2.343	0.7136	≤ .01	0.94	3.74
lgth	-0.016	0.0050	≤ .01	-0.03	-0.01
comv	0.201	0.0234	≤ .01	0.16	0.25
comc2	-2.486	0.6150	≤ .01	-3.69	-1.28
comc3	1.986	0.5716	≤ .01	0.87	3.11
tpfv2	0.343	1.0667	> .05	-1.75	2.43
tpfv3	1.706	1.1916	> .05	-0.63	4.04
		Std Err.	95% Conf. Int.		
phi					
phi_1	1		(constrained)		
phi_2	0.325	0.0397	0.2468	0.4025	
phi_3	0.120	0.0169	0.0868	0.1532	
phi_4	0.022	0.0085	0.0059	0.0393	
phi_5	0		(base outcome)		

Table V shows the value of Wald $\chi2$ (7) = 126.66 and Prob> $\chi2$ = 0.0000. The results indicate that our model is statistically significant compared to the null model, excluding any predictors. Furthermore, the outcomes of the ordered logistic regression model, particularly the column P > | z |, reveal that company-type (*ltdp)*, distance (*lgth)*, value of commodity (*comv)*, and class. of comm (*comc)* are statistically significant, while the type of freight vehicle (*tpfv)* is not.

B. Odd Ratio Analysis

We used the *listcoef* command to obtain the odds ratio (OR) value of the SL regression model. The odd ratio (OR) values obtained are shown in Table VI.

TABLE VI
THE ODD RATIO (OR) VALUES OBTAINED USING STEREOTYPE LOGISTIC (SL) REGRESSION MODEL

Variables	**Coeff. (=b)**	**E^b(= OR)**	**Std. Err.**	**95% Conf. Int.**	
ltdp	2.3426	10.408	0.7136	0.9439	3.7413
lgth	-0.0161	0.984	0.0050	-0.0260	-0.0063
comv	0.2013	1.223	0.0234	0.1554	0.2473
comc2	-2.4858	0.083	0.6150	-3.6912	-1.2803
comc3	1.9860	7.286	0.5716	0.8657	3.1062
tpfv2	0.3428	1.409*	1.0667	-1.7479	2.4336
tpfv3	1.7064	5.509*	1.1916	-0.6292	4.0419

The table presents the odds ratios (ORs, e^b) for five predictor variables compared to the base category, labeled as "Odds of 5 vs. 1" at the top. In the first group, ORs for company type (*ltdp)*, distance (*lgth)*, and comm value. (*comv)* are 10.408, 0.984, and 1.223. In the second group, ORs for *comc2* and *comc3* are 0.083 and 7.286, respectively. In the third group, the ORs for *tpfv2* and *tpfv3* are 1.409 and 5.509, respectively. The variables with the strongest effects are *ltdp* (OR = 10.408) and *comc3* (OR = 7.286), both exhibiting significant positive association. In contrast, *comc2* demonstrates a significant negative association with the outcome (OR = 0.083). Meanwhile, variables such as *tpfv2* and *tpfv3* exhibit weaker or less significant effects based on their confidence intervals.

Category comparison for Y = 5 vs Y = 1, Y = 5 vs Y = 2, Y = 5 vs Y = 3, and Y = 5 vs Y = 4 are provided in Table VII.

The values in this table were computed manually. For instance, for the *ltdp* predictor, the OR of being in the base category 5 is calculated as follows:

- OR (5,1) = e^(1 × 2.3426) = 10.408
- OR (5,2) = e^(0.3246 × 2.3426) = 2.1391
- OR (5,3) = e^(0.120 × 2.3426) = 1.3246
- OR (5,4) = e^(0.0226 × 2.3426) = 1.0544

The values of the variables company type (*ltdp*), distance (*lgth*), value of the commodity (*comv*), and classification of the commodity (*comc2* and *comc3*) can be found in Table VII.

The variable *ltdp* (company type) as a predictor has an OR of 10.4083, meaning that the odds of being in category 5 versus category 1 are 10.4083 times higher for limited partnership companies compared to non-limited partnership companies (smaller companies), assuming all other predictors are held constant. For the variable *lgth* (distance), an OR of

0.9840, which is less than 1, signifies that the odds decrease by a factor of 0.9840 for every one-unit increase in distance, assuming all other predictors remain constant. For the variable *comv* (commodity value), the odds of being in category 5 versus category 1 increase by a factor of 1.223 for each one-unit increase in commodity value.

TABLE VII
COMPARISON OF THE OR OF THE PREDICTOR VARIABLES

Category Comparison	**Y = 5 Vs Y =1**	**Y = 5 Vs Y =2**	**Y = 5 Vs Y =3**	**Y = 5 Vs Y =4**
Variables	*OR*	*OR*	*OR*	*OR*
ltdp	10.4083	2.1391	1.3246	1.0544
lgth	0.9840	0.9948	0.9981	0.9996
comv	1.2230	1.0675	1.0245	1.0046
comc2	0.0833	0.4462	0.7421	0.9454
comc3	7.2863	1.9053	1.2691	1.0459

For the variable *ltdp*, the OR values for category comparisons (category 5 vs categories 1, 2, 3, and 4) are 10.4083, 2.1391, 1.3246, and 1.0544, respectively. Similarly, for the variable *lgth*, the odds decrease across the same category comparisons by factors of 0.9840, 0.9948, 0.9981, and 0.9996 for each one-unit increase in distance. The same interpretation process applies to other predictors, such as *comc2* and *comc3*. Table VII highlights that the five predictor variables exhibit different OR values. A significant relationship exists between the signs of the model coefficients and the OR values of the predictor variables. Among all variables, the company type or scale (*ltdp*) stands out as having the highest OR value.

C. Discussion

The findings from the logistic regression analysis provide critical insights into factors influencing shipment size decisions in cross-border freight transportation. The stereotype logistic regression (SL regression) model demonstrated strong statistical significance, as indicated by a Wald χ^2 value of 126.66 and a p-value of 0.0000. Among the predictor variables, company type (*ltdp*) showed the greatest influence, with an odds ratio (OR) of 10.408, suggesting that larger companies have a substantially higher likelihood of managing high-volume shipments. This finding aligns with the concept of economies of scale, wherein larger firms leverage their financial and operational capacity to optimize freight logistics, as emphasized in previous studies on cross-border transportation efficiency.

The negative association between distance (*lgth*) and shipment size (OR = 0.984) makes a significant contribution to the discourse on transport logistics, indicating that longer transportation distances deter large shipment sizes. This finding corroborates prior research linking extended distances with heightened logistical complexities and costs, often prompting firms to prioritize smaller, more manageable shipments. However, the relatively weak correlation observed in this study highlights the predominant role of other factors, such as company type and commodity value, over transport distance in influencing shipment size decisions.

The positive association between commodity value (*comv*) and shipment size (OR = 1.223) reinforces the strategic preference for transporting high-value commodities in larger quantities to maximize profitability. This aligns with established logistics theories, emphasizing the balance between shipment volume and commodity value as a key determinant in freight optimization. Furthermore, the nuanced impact of commodity classification (*comc*) on shipment size—where *comc3* showing a strong positive correlation (OR = 7.286) and *comc2* demonstrating a negative relationship (OR = 0.083) highlights the need for tailored strategies for different commodity types. These results extend the literature by providing micro-level evidence on how specific commodity characteristics influence shipment size decisions. The findings also address inefficiencies in cross-border freight transportation, such as overloading and underutilization, which contribute to higher operational costs and environmental risks. This study emphasizes the need for strategic interventions, including consolidating low-value commodities into larger shipments and implementing regulatory measures to optimize transport capacity. Such measures align with the sustainability goals highlighted in prior studies and provide actionable recommendations for improving cross-border logistics.

By leveraging micro-level data and employing SL regression, this research contributes to a broader understanding of shipment size determinants, addressing gaps identified in earlier studies. The methodological approach adopted in this study, which uses disaggregated data to model shipment size decisions, offers a novel perspective that complements existing literature, predominantly focused on mode choice and aggregate data analyses. Future research should explore the interplay of these factors with evolving trade policies and infrastructure developments to provide a more comprehensive framework for optimizing cross-border freight transportation.

IV. CONCLUSION

This study analyzes cross-border freight transportation between West Kalimantan, Indonesia, and Sarawak, Malaysia, focusing on shipment size as a determinant of logistical efficiency. The research represents a pioneering effort in leveraging micro-level data and stereotype logistic regression (SL regression) to develop disaggregated strategies for analyzing cross-border freight logistics. Among the five predictor variables examined, vehicle type showed no significant effect, while company type (*ltdp*) and commodity value (*comv*) emerged as significant positive predictors of shipment size. Larger firms demonstrated higher odds ratios (OR) for managing large-volume shipments, particularly in very low (≤ 500 kilograms) and low (501–2,500 kilograms) shipment categories, highlighting the role of greater financial and operational capacity in cross-border logistics.

The positive correlation between commodity value and shipment size suggests that high-value commodities are transported in larger volumes to maximize profitability. A strategic recommendation is provided to consolidate low-value commodities into larger shipments to balance this effect. Interestingly, no significant relationship was observed between shipment size and transport distance (*lgth*), indicating that other factors, such as company characteristics and commodity value, dominate in influencing shipment size decisions.

The findings further emphasize the risks associated with overloading and improper utilization of transport capacity, such as increased road damage and environmental hazards. This emphasizes addressing these inefficiencies through targeted interventions and regulatory improvements. By utilizing stereotype logistic regression, this study contributes to understanding the relationships between shipment size and various predictor variables, offering actionable insights for policymakers and logistics practitioners to enhance the efficiency and sustainability of cross-border freight transportation.

Future research should explore the interaction of these factors with evolving trade policies and infrastructure development, including the implementation of specific regulatory frameworks to promote shipment efficiency or investments in infrastructure that facilitate larger shipment sizes. Additionally, investigating the applicability of stereotype logistic regression in other logistics contexts, such as non-marine commodities or different regional trade corridors, could expand the understanding of its utility. Longitudinal studies analyzing temporal changes in shipment size determinants are also recommended to capture dynamic shifts influenced by global trade trends and technological advancements.

ACKNOWLEDGMENT

The authors express their gratitude to the Fish Quarantine and Inspection Agency for Quality Control and Fishery Products Safety (BKIPM) in Entikong, West Kalimantan, for their invaluable cooperation and for providing access to the necessary data.

REFERENCES

[1] R. Jakab, "Prohibition of cross-border water transport in the conditions of the Slovak Republic and its legal consequences," *J. Agric. Environ. Law*, vol. 18, no. 35, pp. 49–63, Dec. 2023, doi: 10.21029/JAEL.2023.35.49.

[2] H. Wei, "Optimal design of an integrated cross-border logistics network for China's inland regions," *J. Coast. Res.*, vol. 37, no. 3, pp. 644–655, May 2021, doi: 10.2112/jcoastres-d-20-00051.1.

[3] F. Ulengin et al., "A simulation-based approach for improving the largest border crossing between Europe and Turkey," *Transp. Policy*, vol. 114, pp. 350–363, Dec. 2021, doi: 10.1016/j.tranpol.2021.10.011.

[4] G. Daudin, J. Hericourt, and L. Patureau, "International transport costs: New findings from modeling additive costs," *J. Econ. Geogr.*, vol. 22, no. 5, pp. 989–1044, Sept. 2022, doi: 10.1093/jeg/lbac007.

[5] Y. Wang et al., "The volume of trade-induced cross-border freight transportation has doubled and led to 1.14 gigatons CO2 emissions in 2015," *One Earth*, vol. 5, no. 10, pp. 1165–1177, Oct. 2022, doi: 10.1016/j.oneear.2022.09.007.

[6] C. Liu, J. Wu, and H. L. Jayetileke, "Overseas warehouse deployment for cross-border e-commerce in the context of the Belt and Road Initiative," *Sustainability*, vol. 14, no. 15, pp. 1–16, Aug. 2022, doi: 10.3390/su14159642.

[7] B. Dong, M. Duan, and Y. Li, "Exploration of joint optimization and visualization of inventory transportation in agricultural logistics based on ant colony algorithm," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–12, June 2022, doi: 10.1155/2022/2041592.

[8] R. A. Patil, A. D. Patange, and S. S. Pardeshi, "International transportation mode selection through total logistics cost-based intelligent approach," *Logistics*, vol. 7, no. 3, pp. 1–26, Sep. 2023, doi: 10.3390/logistics7030060.

[9] C. F. Gao, Z. H. Hu, and Y. Z. Wang, "Optimizing the hub-and-spoke network with drone-based traveling salesman problem," *Drones*, vol. 7, no. 1, pp. 1–25, Dec. 2023, doi: 10.3390/drones7010006.

[10] L. Ni and X. Wang, "Load factors of less-than-truckload delivery tours: An analysis with operation data," *Transp. Res. E: Logist. Transp. Rev.*, vol. 150, pp. 1–18, Jun. 2021, doi: 10.1016/j.tre.2021.102296.

[11] E. A. Thompson, R. Abudu, and S. Zheng, "Empirical analysis of multiple-criteria decision-making (MCDM) process for freight transportation mode selection," *J. Transp. Technol.*, vol. 12, no. 1, pp. 28–41, Jan. 2022, doi: 10.4236/jtts.121002.

[12] Y. Lu and S. Wang, "Optimization of joint decision of transport mode and path in multi-mode freight transportation network," *Sensors*, vol. 22, no. 13, pp. 1–18, Jun. 2022, doi: 10.3390/s22134887.

[13] G. Günay, "Shipment size and vehicle choice modeling for road freight transport: A geographical perspective," *Transp. Res. A: Policy Pract.*, vol. 173, p. 103732, Jul. 2023, doi: 10.1016/j.tra.2023.103732.

[14] G. Bray and D. Cebon, "Selection of vehicle size and extent of multi-drop deliveries for autonomous goods vehicles: An assessment of potential for change," *Transp. Res. E: Logist. Transp. Rev.*, vol. 164, p. 102806, Aug. 2022, doi: 10.1016/j.tre.2022.102806.

[15] X. Yang, "Patterns of export shipments," *J. Econ. Stud.*, vol. 50, no. 2, pp. 283–299, Feb. 2023, doi: 10.1108/JES-11-2021-0575.

[16] A. Gohari et al., "Comparative analysis of single- and multi-criteria container transport modes in Peninsular Malaysia," *Int. J. Sustain. Eng.*, vol. 14, no. 5, pp. 1239–1250, Jun. 2021, doi: 10.1080/19397038.2020.1774819.

[17] BKIPM, "Cross-border goods vehicle journey data, marine commodities destined for Malaysia, 2016–2018," Entikong, West Kalimantan, Indonesia, 2018, unpublished report.

[18] U. M. Mbanaso, L. Abrahams, and K. C. Okafor, "Research philosophy, design and methodology," in *Research Techniques for Computer Science, Information Systems and Cybersecurity*, 1st ed., Springer Nature Switzerland AG, May 2023, pp. 1–32, doi: 10.1007/978-3-031-30031-8_6.

[19] A. K. Reda et al., "Temporal stability of shipment size decisions related to choice of truck type," *Transportmetrica A: Transp. Sci.*, vol. 20, no. 3, pp. 1–29, May 2023, doi: 10.1080/23249935.2023.2214635.

[20] Y. Yang et al., "Identifying intercity freight trip ends of heavy trucks from GPS data," *Transp. Res. E: Logist. Transp. Rev.*, vol. 157, pp. 1–21, Jan. 2022, doi: 10.1016/j.tre.2021.102590.

[21] U. Ahmed and M. J. Roorda, "Joint and sequential models for freight vehicle type and shipment size choice," *Transportation*, pp. 1613–1629, Oct. 2023, doi: 10.1007/s11116-022-10289-6.

[22] J. Holguín-Veras, L. Kalahasthi, S. Campbell, C. A. González-Calderón, and X. (Cara) Wang, "Freight mode choice: Results from a nationwide qualitative and quantitative research effort," *Transp. Res. A: Policy Pract.*, vol. 143, pp. 78–120, Jan. 2021, doi: 10.1016/j.tra.2020.11.016.

[23] K. Backhaus, B. Erichson, S. Gensler, R. Weiber, and T. Weiber, *Multivariate Analysis: An Application-Oriented Introduction*, 2nd ed. Springer Gabler, pp. 279–299, Nov. 2022, doi: 10.1007/978-3-658-40411-6.

[24] Statistical Consulting Group, "Logistic regression with STATA," 2021. [Online]. Available: https://stats.oarc.ucla.edu/stata/seminars/stata-logistic/ (accessed Apr. 22, 2021).

[25] J. S. Long and J. Freese, *Regression Models for Categorical Dependent Variables Using Stata*, 3rd ed. Texas: Stata Press, 2014, pp. 120–132.

[26] N. M. Zafri, A. Khan, S. Jamal, and B. M. Alam, "Risk perceptions of COVID-19 transmission in different travel modes," *Transp. Res. Interdiscip. Perspect.*, vol. 13, pp. 1–10, Jan. 2022, doi: 10.1016/j.trip.2022.100548.

[27] J. Romão and Y. Bi, "Determinants of collective transport mode choice and its impacts on trip satisfaction in urban tourism," *J. Transp. Geogr.*, vol. 94, pp. 1–9, Jun. 2021, doi: 10.1016/j.jtrangeo.2021.103094.

[28] A. Nikitas, A. E. Vitel, and C. Cotet, "Autonomous vehicles and employment: An urban futures revolution or catastrophe?," *Cities*, vol. 114, pp. 1–14, Jul. 2021, doi: 10.1016/j.cities.2021.103203.

[29] A. B. A. Al-Mekhlafi et al., "Impact of safety culture implementation on driving performance among oil and gas tanker drivers: A partial least squares structural equation modelling (PLS-SEM) approach," *Sustainability*, vol. 13, no. 16, pp. 1–17, Aug. 2021, doi: 10.3390/su13168886.

[30] A. Allee, L. R. Lynd, and V. Vaze, "Cross-national analysis of food security drivers: Comparing results based on the Food Insecurity Experience Scale and Global Food Security Index," *Food Secur.*, vol. 13, no. 5, pp. 1245–1261, Mar. 2021, doi: 10.1007/s12571-021-01156-w.

[31] P. Thaithatkul, S. Chalermpong, W. Laosinwattana, and H. Kato, "Mobility, activities, and happiness in old age: Case of the elderly in Bangkok," *Case Stud. Transp. Policy*, vol. 10, no. 2, pp. 1462–1471, Jun. 2022, doi: 10.1016/j.cstp.2022.05.010.

[32] I. E. Ceyisakar et al., “Ordinal outcome analysis improves the detection of between-hospital differences in outcome,” *BMC Med. Res. Methodol.*, vol. 21, no. 1, pp. 1–11, Jan. 2021, doi: 10.1186/s12874-020-01185-7.

[33] X. Liu, *Applied Ordinal Logistic Regression Using Stata*, 1st ed. California: Sage Publications, Inc., 2016, pp. 235–357.

[34] M. E. Lelisho, A. A. Wogi, and S. A. Tareke, “Ordinal logistic regression analysis in determining factors associated with socioeconomic status of household in Tepi Town, Southwest Ethiopia,” *Sci. World J.*, vol. 2022, pp. 1–9, Feb. 2022, doi: 10.1155/2022/2415692.

[35] R. Williams, “Generalized ordered logit/partial proportional odds models for ordinal dependent variables,” *Stata J.*, vol. 6, no. 1, pp. 58–82, 2006, doi: 10.1177/1536867X0600600104.

[36] R. Williams, “Gologit2: A program for generalized logistic regression/partial proportional odds models for ordinal dependent variables,” Notre Dame, United States, 2005, pp. 1–18.

[37] R. Williams, “Ordered logit models—Basic & intermediate topics,” University of Notre Dame, Indiana, Jan. 2022.

[38] X. Liu, “Fitting stereotype logistic regression models for ordinal response variables in educational research (Stata),” *J. Mod. Appl. Stat. Methods*, vol. 13, no. 2, pp. 528–545, Nov. 2014, doi: 10.22237/jmasm/1414816200.