*Journal of Economics, Entrepreneurship, Management Business and Accounting*, 2026, *4*(5),*1286-1298.*
https://doi.org/10.61255/jeemba.v4i5.1658

# Marketing Mix Factors Associated with Online Purchasing Behaviour for Printed Books: Evidence from Actual Transaction Data

Erick Wahyudyono[1*], Ujang Sumarwan[2], and Arief Safari[3]

School of Business, IPB University, Bogor, 16151, Indonesia

**ABSTRACT**

**Purpose** – This study examines the adjusted associations between recorded transaction attributes and realised online printed-book invoice value at PT XYZ during 2023–2025, using the 7P marketing mix as a conceptual organising lens.
**Design/methodology/approach** – A retrospective observational design was applied to a census of 5,989 eligible completed online printed-book invoices. The unit of analysis was an invoice rather than a unique consumer. Ordinary least squares regression was estimated for logged invoice value with HC3 heteroskedasticity-consistent robust standard errors. The primary explanatory variables were academic-book share, destination region, purchase channel, and transaction year. Average selling price and basket size were excluded from the primary model because they are mechanically related to invoice value.
**Finding/Results –** The model explained 28.4% of the variation in logged invoice value, and the predictors were jointly associated with the outcome. A 10-percentage-point increase in academic-book share was associated with an estimated 3.2% higher invoice value. Invoices delivered to Java outside Jabodetabek and outside Java had estimated values approximately 9.0% and 12.9% higher, respectively, than invoices delivered within Jabodetabek. Marketplace and WhatsApp transactions had estimated values approximately 5.3% and 7.7% higher than webstore transactions. Estimated invoice values were also higher in 2024 and 2025 than in 2023. Robustness checks showed that the estimates were stable after excluding non-individual and influential invoices.
**Originality/Value** – This study provides transaction-level evidence from actual completed online book invoices rather than survey-based perceptions. The recorded attributes are interpreted as observable transaction characteristics conceptually related to the 7P framework, not as direct measures of consumers' perceptions or causal effects of marketing activities.

**ARTICLE INFO**

**Keywords:**
Actual transaction data, consumer purchasing behavior, Indonesia publishing industry, marketing mix,printed books.

**Article Information:**
Received: 15/7/2026
Revise: 29/7/2026
Accepted: 20/8/2026

**ISSN:**
2985-3168 (Online)
2985-3222 (Print)

*Corresponding Author at:
School of Business, IPB University, Jalan Raya Padjajaran, Bogor, 16151, Indonesia.
E-mail address: erickwahyudyono@apps.ipb.ac.id (Erick Wahyudyono)

## 1. Introduction

The creative industry is an important contributor to Indonesia's economic development. According to the Ministry of Creative Economy (2025), the creative economy contributed more than IDR 1,500 trillion to Indonesia's gross domestic product in 2024 and supported more than 26.5 million jobs. Publishing is one of Indonesia's creative-economy subsectors (Jewell, 2019), while Presidential Regulation No. 96 of 2019 concerning the Ministry of Tourism and Creative Economy. In addition to its economic contribution, publishing supports the dissemination of knowledge, educational resources, and cultural products. The continuing production of printed and digital titles, reflected in annual ISBN and e-ISBN records published by the National Library of Indonesia from 2019 to 2024, indicates that the publishing industry has a substantial foundation for continued development. Within this context, online sales have become increasingly relevant for publishers seeking to expand the distribution and accessibility of printed books across geographically dispersed markets.

The expansion of e-commerce has created new opportunities for publishers to reach customers through digital channels. E-commerce enables transactions to be completed across locations and devices, while providing consumers with more convenient and accessible purchasing options (Laudon & Traver 2021). In Indonesia, the growth of online shopping has been supported by technological development and consumers' preference for more efficient purchasing processes (Nasution & Putri 2019). The COVID-19 pandemic also strengthened the role of online channels in the book market, with evidence from Romania showing a positive association between the pandemic period and online book purchases (Dinca & Dima 2022). However, an expansion in the volume or aggregate value of e-commerce does not imply that all completed book transactions generate similar invoice values. Individual invoices may differ according to the composition of purchased books, the destination of the order, the channel through which the transaction is completed, and other recorded characteristics of the transaction.

Most previous studies examining marketing-related factors and purchase outcomes have relied on cross-sectional surveys, consumer perceptions, purchase intentions, or self-reported purchase decisions (Firmansyah et al. 2021; Pasharibu et al. 2020; Suyanto & Dewi, 2023; Widyastuti et al. 2020). These approaches are valuable for explaining how consumers evaluate marketing activities and form purchase-related intentions. Nevertheless, perceptions and intentions do not necessarily correspond to the monetary value eventually recorded in a completed transaction (Morwitz, 2014). Conversely, transaction records provide evidence of realised purchases, but they do not reveal consumers' underlying motivations, perceptions of marketing activities, promotional exposure, non-purchase decisions, or the complete decision-making process. Thus, survey-based and transaction-based studies provide complementary, rather than interchangeable, forms of evidence (Alberini, 2019).

This distinction reveals an empirical and conceptual gap. Empirical evidence based on administrative invoice records from Indonesian online printed-book retailers remains limited, particularly with respect to variation in the monetary value of completed invoices. The conceptual issue is equally important. Administrative fields such as product composition, purchase channel, destination, service process, and order fulfilment are observable transaction characteristics; they are not psychometric indicators of consumers' perceived product, price, place, promotion, people, process, or physical evidence. Treating such fields as direct measures of latent 7P constructs would conflate operational records with the measurement of theoretical constructs (Baumgartner & Weijters, 2019; Jarvis et al., 2003). Accordingly, the 7P

framework is used only as a broad organising lens for interpreting transaction-related information (Rafiq & Ahmed, 1995), rather than as a set of directly measured marketing constructs or a basis for causal claims.

Against this empirical and conceptual gap, this study examines 5,989 completed online printed-book invoices recorded by PT XYZ between 2023 and 2025. The unit of analysis is an individual invoice rather than a unique consumer, and the outcome is the logged monetary value of each completed invoice. The study investigates the associations between recorded transaction attributes and variation in online printed-book invoice value while recognising the observational nature and limitations of the available data. The novelty of this study lies not in claiming to test the 7P marketing mix as directly measured consumer constructs, but in repositioning recorded administrative fields as observable transaction attributes and using them to complement perception-based marketing research. This approach provides transaction-level evidence on realised invoice values in the online printed-book market without claiming to explain consumer purchasing behaviour comprehensively or to establish causal effects.

## 2. Methodology

This study employed a retrospective quantitative observational design using administrative transaction records from PT XYZ, which were originally created for operational purposes and subsequently analysed for research (Connelly et al., 2016; Hand, 2018). The research process was conducted from November 2025 to April 2026. PT XYZ was purposively selected because it provided access to completed online printed-book transaction records from customers located across Indonesia. The unit of analysis was an individual completed invoice rather than a unique consumer; therefore, more than one invoice could potentially originate from the same customer. The study applied a census approach by including all 5,989 eligible completed online printed-book invoices recorded between 2023 and 2025. Accordingly, no sampling was conducted at the transaction-record level. The dataset captures realised transactions recorded by the company but does not contain consumers' motivations, perceptions of marketing activities, promotional exposure, non-purchase decisions, or the complete consumer decision-making process.

The 7P marketing mix was used only as a broad conceptual lens for organising the transaction-related information available in PT XYZ's database. This framework combines the conventional marketing-mix elements with people, process, and physical evidence (Kotler & Armstrong, 2018; Booms & Bitner, 1981). However, the recorded fields were not treated as direct psychometric measures of consumers' perceptions of the 7P constructs. Rather, they were interpreted as observable transaction attributes. The dependent variable was the natural logarithm of the monetary value recorded for each completed online printed-book invoice. The primary explanatory variables were academic-book share, destination region, and purchase channel, while transaction year was included as a temporal control. The operational definitions, measurements, and analytical roles of all recorded variables are presented in Table 1.

**Table 1.** Operational Definitions and Measurements of Research Variables

| Variable | Operational Definition | Measurement | Scale | Data Source/Conceptual Background |
|---|---|---|---|---|
| Logged online printed-book invoice value | Monetary value recorded for each completed online printed-book invoice | Natural logarithm of invoice value, Ln (Invoice Value) | Ratio | PT XYZ administrative transaction database |
| Academic-book share | Composition of academic books within an individual invoice | Proportion of academic books among the total books in an invoice (0–1) | Ratio | Kotler & Armstrong (2018) |
| Destination region | Recorded destination of the completed order | Jabodetabek as the reference category; Java outside Jabodetabek; outside Java | Dummy | Kotler & Armstrong (2018) |
| Purchase channel | Channel through which the completed order was recorded | Webstore as the reference category; marketplace; WhatsApp | Dummy | Kotler & Armstrong (2018); Laudon & Traver (2021) |
| Transaction year | Year in which the invoice was completed | 2023 as the reference category; 2024; 2025 | Dummy | PT XYZ administrative transaction database; temporal control |
| Average selling price | Average selling price per book within an invoice | Natural logarithm of average selling price | Ratio, log-transformed | Kotler and Armstrong (2018) |
| Basket size | Number of books purchased in an individual invoice | Number of books per invoice | Ratio | Booms & Bitner (1981); Martin et al. (2020); Wirtz & Lovelock (2022) |
| Service response time | Recorded duration of service response for a completed order | 1 day as the reference category; 2 days; more than 2 days | Dummy | Booms & Bitner (1981); Grosse et al. (2017); Wirtz & Lovelock (2022) |
| Packaging type | Packaging used for the completed order | Paper as the reference category; cardboard | Dummy | Booms & Bitner (1981); Escursell et al. (2021) |

***Note.*** *Academic-book share, destination region, and purchase channel were included in the primary model, with transaction year as a temporal control. Average price, basket size, service response time, and packaging type were reported descriptively only.*

Academic-book share represents the composition of books within an invoice and should not be interpreted as a direct measure of overall product quality or product strategy. Similarly, destination region represents the recorded fulfilment location, whereas purchase channel represents the route through which a completed order was recorded. In the context of e-commerce, the channel through which an order is completed does not directly measure consumers' exposure to promotional activities (Laudon & Traver, 2021). Average selling price and basket size were excluded from the primary regression model because both variables are mechanically related to invoice value. The monetary value of an invoice is directly influenced by the price of purchased books and the number of books included in the order. Including these variables as primary predictors would therefore create an accounting decomposition of invoice value rather than an independent assessment of transaction attributes associated with invoice value. Service response time and packaging type were also excluded from the primary model because these operational characteristics may be determined after the scale and monetary

value of an order are already known. They were retained for descriptive reporting only and were not interpreted as explanatory drivers of invoice value.

Descriptive statistics were first used to summarise the characteristics of the 5,989 completed invoices. Continuous variables were described using appropriate measures of central tendency and dispersion, whereas categorical variables were reported as frequencies and percentages. The primary analysis used multiple linear regression estimated by Ordinary Least Squares (OLS) to examine conditional associations between the recorded transaction attributes and logged invoice value (Wooldridge, 2020). The model was specified as follows:

$$\ln(Y_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \varepsilon_i$$

Where:

| | | | |
|---|---|---|---|
| $Y_i$ | : Online printed-book invoice value | $X_{4i}$ | : Marketplace |
| $X_{1i}$ | : Academic-book share | $X_{5i}$ | : WhatsApp |
| $X_{2i}$ | : Java outside Jabodetabek | $X_{6i}$ | : Transaction year 2024 |
| $X_{3i}$ | : Outside Java | $X_{7i}$ | : Transaction year 2025 |

In this model, Jabodetabek, webstore, and 2023 were used as the respective reference categories. Because academic-book share was coded as a proportion from 0 to 1, its coefficient represents the adjusted association corresponding to a one-unit change in the share. For practical interpretation, a 10-percentage-point increase in academic-book share corresponds to a 0.10-unit increase in the variable. The coefficients for destination region and purchase channel compare the expected logged invoice value of invoices in each category with that of the relevant reference category, after accounting for the other transaction attributes included in the model. The transaction-year variables capture temporal differences in invoice values between 2024 or 2025 and the 2023 reference period. Because the unit of analysis is an individual invoice rather than a unique consumer, the estimates should be interpreted at the transaction level. Accordingly, the model identifies adjusted associations between recorded transaction attributes and invoice value, but it does not estimate consumer preferences, purchase probability, demand elasticity, or causal effects of marketing activities.

Regression diagnostics included variance inflation factors to assess multicollinearity, residual-versus-fitted plots to assess linearity and variance patterns, and normal probability plots of residuals. Heteroskedasticity was formally assessed using the Breusch-Pagan test. In line with regression diagnostic procedures (Hair et al., 2019), the primary statistical inferences were based on heteroskedasticity-consistent HC3 robust standard errors, including robust confidence intervals and *p*-values. Influence diagnostics were conducted using leverage values and Cook's distance. As robustness checks, the primary model was re-estimated after excluding the 47 invoices identified as non-individual transactions and after excluding observations identified as unduly influential. An autocorrelation test was not applied because the data consisted of individual invoices rather than an ordered time-series dataset. Because the dataset represents a census of eligible company transactions, inferential statistics were used to quantify uncertainty in the estimated associations within the observed transaction-generating process, rather than to claim representativeness of all Indonesian book consumers.

## 3. Result and Discussion

### 3.1. Transaction Characteristics

This section describes the characteristics of the 5,989 completed online printed-book invoices recorded by PT XYZ between 2023 and 2025. The descriptive analysis covers invoice type,

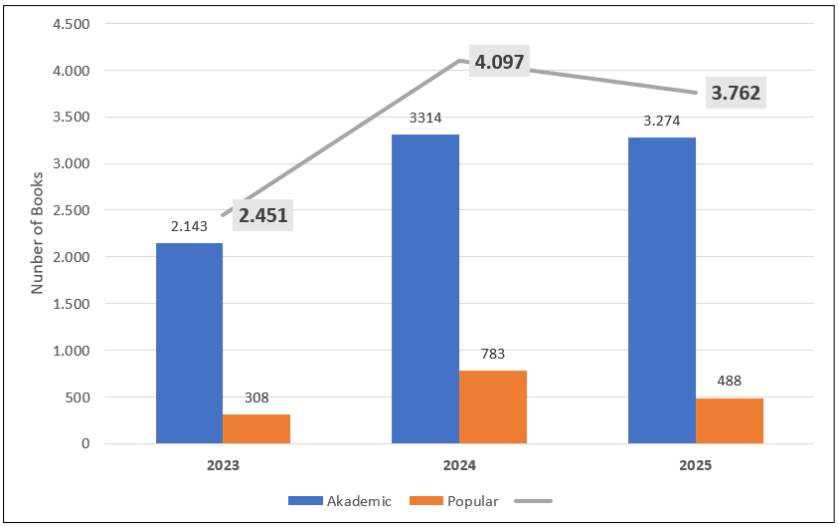
purchase channel, destination region, books per invoice, packaging type, and invoice-value category.

**Table 2.** Transaction Characteristics of Completed Online Printed-Book Invoices

| Characteristic | Category | *n* | Percentage (%) |
|---|---|---|---|
| Invoice type | Individual | 5.942 | 99,22 |
| | Non-Individual | 47 | 0,78 |
| Purchase channel | Webstore | 1.633 | 27,27 |
| | Marketplace | 3.580 | 59,78 |
| | WhatsApp | 776 | 12,96 |
| Destination region | Jabodetabek | 2.222 | 37,10 |
| | Java Non Jabodetabek | 1.860 | 31,06 |
| | Outside Java | 1.907 | 31,84 |
| Books per invoice | Fewer than 3 books | 5.467 | 91,28 |
| | 3 – 5 books | 366 | 6,11 |
| | More than 5 books | 156 | 2,61 |
| Packaging type | Paper | 5.936 | 99,12 |
| | Cardboard/Box | 53 | 0,88 |
| Invoice value | < IDR 100.000 | 2.152 | 35,93 |
| | IDR 100.000 – Rp 200.000 | 2.788 | 46,55 |
| | > IDR 200.000 | 1.049 | 17,52 |

*Source: Total number of completed invoices = 5,989.*

Table 2 indicates that the dataset was dominated by invoices classified as individual transactions, which accounted for 99.22% of all completed invoices, whereas non-individual transactions represented only 0.78%. Marketplace was the most frequently recorded purchase channel, accounting for 59.78% of invoices, followed by webstore and WhatsApp. The destination distribution was relatively balanced across Jabodetabek, Java outside Jabodetabek, and regions outside Java, with each category representing approximately one-third of the observed invoices. Most invoices contained fewer than three books (91.28%), and paper packaging was used in nearly all transactions. In terms of invoice value, the IDR 100,000–200,000 category represented the largest proportion of invoices, accounting for 46.55% of the dataset.




*Source: PT XYZ, data processed (2026)*

**Figure 1.** Number of printed books purchased online during 2023–2025

Figure 1 shows that academic printed books consistently accounted for a larger number of books purchased online than popular books throughout 2023–2025. Total annual purchase volume varied across the observation period and reached its highest level in 2024, at 4,097 copies. The figure describes the observed annual pattern but does not identify the external, seasonal, or marketing factors underlying the year-to-year variation.

### 3.2. Multiple Linear Regression Analysis

A primary OLS regression model was estimated using all 5,989 completed invoices to examine the adjusted associations between recorded transaction attributes and online printed-book invoice value. The model included academic-book share, destination region, purchase channel, and transaction-year controls. Table 3 reports the coefficient estimates, HC3 heteroskedasticity-consistent robust standard errors, 95% confidence intervals, and *p*-values. Jabodetabek, webstore, and 2023 were used as the respective reference categories.

**Table 3.** Primary Regression Results for Logged Online Printed-Book Invoice Value

| Predictor | Coefficient (B) | HC3 Robust SE | 95% CI | *p*-value |
|---|---|---|---|---|
| Constant | 10.284 | 0.041 | [10.204, 10.364] | < .001 |
| Academic-book share | 0.318 | 0.047 | [0.226, 0.410] | < .001 |
| Java outside Jabodetabek | 0.086 | 0.019 | [0.049, 0.123] | < .001 |
| Outside Java | 0.121 | 0.018 | [0.086, 0.156] | < .001 |
| Marketplace | 0.052 | 0.017 | [0.019, 0.085] | .002 |
| WhatsApp | 0.074 | 0.021 | [0.033, 0.115] | < .001 |
| Transaction year 2024 | 0.067 | 0.018 | [0.032, 0.102] | < .001 |
| Transaction year 2025 | 0.115 | 0.020 | [0.076, 0.154] | < .001 |
| Model statistics | | | | |
| Number of invoices ((N)) | 5,989 | | | |
| $R^2$ | 0.284 | | | |
| Adjusted $R^2$ | 0.283 | | | |
| Robust Wald (F (7,5,981)) | 337.62 | | | < .001 |

*Note. HC3 = robust standard error.*

The robust joint-model test indicated that the predictors were jointly associated with logged invoice value, and the primary model accounted for 28.4% of the variation in logged invoice value, with an adjusted (R^2) of 0.283. All estimated associations in the primary model remained statistically different from zero based on HC3 robust standard errors. Academic-book share was positively associated with invoice value; a 10-percentage-point increase in academic-book share was associated with an estimated 3.2% higher invoice value. Both destination categories beyond Jabodetabek also showed higher estimated invoice values, with the association for invoices sent outside Java larger than that for invoices sent to Java outside Jabodetabek.

The purchase-channel estimates indicated higher invoice values for both marketplace and WhatsApp transactions relative to webstore transactions. The temporal controls also showed that estimated invoice values were higher in 2024 and 2025 than in 2023, with the larger difference observed in 2025. These results describe adjusted differences across the recorded transaction attributes and periods; they do not indicate causal effects.

### 3.3. Regression Diagnostic Results

Diagnostic tests were conducted to evaluate the adequacy of the primary regression estimates reported in Table 3. The assessment examined m
ulticollinearity among the explanatory variables, residual variance and distribution patterns, and the potential influence of individual invoices on the estimated coefficients. The diagnostic results are summarised in Table 4.

**Table 4.** Regression Diagnostic Results for the Primary Model

| Diagnostic | Result | Interpretation |
|---|---|---|
| Variance inflation factor range | 1.08–1.42 | No multicollinearity concern |
| Highest variance inflation factor | 1.42 | Acceptable |
| Breusch–Pagan test statistic | 86.47 | Heteroskedasticity detected |
| Breusch–Pagan test *p*-value | < .001 | Evidence of non-constant residual variance |
| Residual-versus-fitted plot | Mild funnel-shaped pattern | Consistent with heteroskedasticity |
| Normal probability plot of residuals | Approximately linear with minor tail deviations | No substantial departure from normality |
| Maximum Cook's distance | 0.081 | No extremely influential observation |
| Invoices identified for sensitivity analysis | 14 | Examined in the robustness analysis |

The VIF values were low, indicating no problematic multicollinearity among the explanatory variables. In contrast, the Breusch–Pagan test and the residual-versus-fitted plot indicated heteroskedasticity, supporting the use of HC3 robust standard errors for statistical inference. The normal probability plot showed only minor deviations in the residual tails. In addition, the maximum Cook's distance was 0.081, indicating that no individual invoice exerted an extreme influence on the primary model. Nevertheless, 14 invoices identified through the influence diagnostics were examined further in the robustness analysis.

### 3.4. Robustness Checks

Robustness analyses were conducted to examine whether the primary associations were sensitive to the inclusion of non-individual transactions and potentially influential invoices. The regression model was re-estimated after excluding 47 non-individual invoices and, separately, after excluding the 14 invoices identified through the influence diagnostics. The results are presented in Table 5.

**Table 5.** Robustness Checks for the Primary Regression Model

| Predictor | Primary model: all invoices (N = 5,989) | Excluding non-individual invoices (N = 5,942) | Excluding influential invoices (N = 5,975) |
|---|---|---|---|
| Academic-book share | 0.318 (0.047) | 0.321 (0.047) | 0.316 (0.046) |
| Java outside Jabodetabek | 0.086 (0.019) | 0.087 (0.019) | 0.085 (0.019) |
| Outside Java | 0.121 (0.018) | 0.122 (0.018) | 0.120 (0.018) |
| Marketplace | 0.052 (0.017) | 0.051 (0.017) | 0.053 (0.017) |
| WhatsApp | 0.074 (0.021) | 0.073 (0.021) | 0.075 (0.021) |
| Transaction year 2024 | 0.067 (0.018) | 0.068 (0.018) | 0.066 (0.018) |
| Transaction year 2025 | 0.115 (0.020) | 0.116 (0.020) | 0.114 (0.020) |
| $R^2$ | 0.284 | 0.286 | 0.282 |
| Adjusted $R^2$ | 0.283 | 0.285 | 0.281 |

*Note. HC3 robust standard errors are shown in parentheses.*

The estimated coefficients remained highly stable across the three model specifications. Excluding non-individual invoices and excluding influential invoices did not change the direction of any association or produce meaningful changes in coefficient magnitude. Model fit also remained similar, with adjusted $R^2$ values ranging from 0.281 to 0.285. These results indicate that the primary findings were not driven by the small number of non-individual transactions or by the invoices identified through the influence diagnostics.

**3.5. Discussion**

The positive association between academic-book share and invoice value should be interpreted as a transaction-composition pattern rather than evidence that academic books are inherently superior or that consumers necessarily prefer them. Because academic-book share was coded from 0 to 1, a 10-percentage-point increase in the academic-book share of an invoice was associated with an estimated 3.2% higher invoice value. This pattern may reflect the tendency for invoices with a larger academic-book component to involve more specialised book selections. Evidence from online bookselling indicates that access to a broader range of titles can create substantial value for consumers, particularly where online channels make less commonly available titles accessible (Brynjolfsson et al., 2023). Nevertheless, the present measure captures the composition of books within an invoice, not product quality, consumer motives, or perceived relevance. The observed association therefore cannot be interpreted as a direct effect of product strategy on consumer behaviour.

The destination-region estimates indicate that invoices delivered to Java outside Jabodetabek and to regions outside Java had higher estimated values than invoices delivered within Jabodetabek. The larger association for regions outside Java indicates that more distant destinations were not necessarily associated with lower-value completed transactions. This interpretation is consistent with evidence from online book retail showing that consumers' location and the availability of local retail alternatives can shape online purchasing patterns (Forman et al., 2009). Research also shows that reduced transportation costs are an important source of value from e-commerce, although the magnitude of this benefit may vary across products and locations (Huang & Bronnenberg, 2023). However, destination region in the present study represents the recorded fulfilment location rather than a direct measure of PT XYZ's distribution strategy, local book availability, delivery costs, or consumer access constraints. The finding should consequently be interpreted as an adjusted geographical difference in realised invoice value, not as evidence that geographical distance itself increases purchasing value.

Both marketplace and WhatsApp transactions were associated with higher estimated invoice values than webstore transactions. These differences should not be interpreted as evidence of greater promotional effectiveness, because purchase channel identifies the route through which an order was completed rather than consumers' exposure to promotional activities. Previous research shows that channel-related purchasing patterns can differ across product categories and customer segments (Kondo & Okubo, 2022). Earlier transaction-based research also found that customers using multiple transaction channels can differ in revenue and purchasing profiles from other customers (Kumar & Venkatesan, 2005). However, PT XYZ's database does not identify whether the same customer used more than one channel, nor does it record customer acquisition, promotional exposure, or loyalty status. Channel-related differences in the present study may therefore reflect variation in order composition, customer profiles, or the types of books commonly purchased through each channel.

The year controls showed higher estimated invoice values in both 2024 and 2025 relative to 2023, with the larger difference observed in 2025. This pattern is noteworthy because the total number of books purchased reached its highest level in 2024, whereas the adjusted estimated invoice value was higher in 2025. The difference suggests that annual changes in invoice value may not correspond directly to changes in the number of books purchased. However, the model was not designed to identify the sources of this temporal variation, such as changes in book composition, product availability, price structure, or broader market conditions. Transaction year was therefore included to account for temporal differences rather than to estimate the effect of a particular annual event or marketing activity.

From a practical perspective, the findings support the use of transaction data to monitor academic-book composition, regional order profiles, and completed-order channels. PT XYZ may use these patterns to identify transaction profiles that warrant further commercial analysis. At the same time, the model explained 28.4% of the variation in logged invoice value, indicating that a substantial proportion of variation was associated with factors not recorded in the database. These may include customer characteristics, book titles, promotional exposure, shipping charges, acquisition costs, and order profitability. In addition, the unit of analysis was an invoice rather than a unique consumer, meaning that multiple invoices may have originated from the same customer. Decisions concerning product promotion, regional expansion, or channel investment should therefore be supported by additional customer-level and operational data. The reported associations are observational and transaction-based and do not establish causal effects of marketing actions on consumer purchasing behaviour.

## 4. Conclusion and Suggestion

This study examined the conditional associations between observable transaction attributes and logged online printed-book invoice value using a census of 5,989 completed invoices recorded by PT XYZ from 2023 to 2025. The results showed that academic-book share, destination region, purchase channel, and transaction year were each associated with variation in invoice value after adjustment for the remaining recorded attributes. Invoices with a higher academic-book share, those delivered outside Jabodetabek, and those completed through marketplace or WhatsApp had higher estimated values than their respective reference categories. Estimated invoice values were also higher in 2024 and 2025 than in 2023. These findings describe realised transaction patterns within PT XYZ's database; they do not demonstrate consumer preferences, demand elasticities, or causal effects of marketing activities.

The findings suggest that PT XYZ may use transaction records to monitor academic-book composition, regional order profiles, and channel-specific order patterns. Such monitoring can support more targeted inventory planning, regional market assessment, and channel evaluation. However, decisions concerning promotion, channel investment, and geographic expansion should also consider customer acquisition costs, delivery costs, margins, repeat purchases, and promotional exposure. Transaction-based analysis should therefore complement, rather than replace, customer-perception research and other evidence required for strategic marketing decisions.

## 5. Limitations and Future Research

This study has several limitations. First, the administrative transaction database recorded completed invoices rather than consumers' motivations, perceptions, promotional exposure, product evaluations, or decisions not to purchase. The observed variables were therefore transaction-based proxies and did not directly measure the latent dimensions of the 7P marketing mix. Second, the unit of analysis was an invoice rather than a unique customer, so multiple invoices may have originated from the same customer. Third, average selling price and basket size were intentionally excluded from the primary model because they are mechanically related to invoice value, while service response time and packaging type were not treated as explanatory variables because they may be determined after the transaction scale is known. Finally, the model explained 28.4% of the variation in logged invoice value and was based on records from a single company; thus, the findings may not generalise to all Indonesian book consumers or online book retailers.

Future research could link transaction records with customer-level information, product-title data, promotional exposure, delivery costs, margins, and repeat-purchase histories. Combining transaction data with survey measures would also enable researchers to examine how actual purchases relate to consumer motivations, perceived product value, service evaluations, and channel preferences. Studies using longer time periods, monthly or quarterly controls, multiple firms, and customer-level repeated observations would provide stronger evidence on temporal patterns and external validity. Where feasible, field experiments or quasi-experimental designs could evaluate the causal effects of specific promotional, channel, or merchandising interventions on transaction outcomes.

## Declaration of AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT to assist with language refinement and the organisation of the manuscript. After using this tool, the authors reviewed and edited the content as necessary and take full responsibility for the content of the publication.

## Reference

Alberini, A. (2019). Revealed versus stated preferences: What have we learned about valuation and behavior? *Review of Environmental Economics and Policy, 13*(2), 283–298. https://doi.org/10.1093/reep/rez010

Baumgartner, H., & Weijters, B. (2019). Measurement in marketing. *Foundations and Trends in Marketing, 12*(4), 278–400. https://doi.org/10.1561/1700000058

Booms, B. H., & Bitner, M. J. (1981). Marketing strategies and organizational structures for service firms. In J. H. Donnelly & W. R. George (Eds.), *Marketing of services* (pp. 47–51). American Marketing Association.

Brynjolfsson, E., Hu, Y. J., & Smith, M. D. (2023). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science, 49*(11), 1580–1596. https://doi.org/10.1287/mnsc.49.11.1580.20580

Connelly, R., Playford, C. J., Gayle, V., & Dibben, C. (2016). The role of administrative data in the big data revolution in social science research. *Social Science Research, 59*, 1–12. https://doi.org/10.1016/j.ssresearch.2016.04.015

Dinca, V. M., & Dima, A.M. (2022). The impact of the covid-19 pandemic on online consumer behavior for the book market in Romania. *E&M: Ekonomics and Management, 25*(4), 155-163. https://doi.org/10.15240/tul/001/2022-4-010

Escursell, S., Llorach-Massana, P., & Roncero, M. B. (2021). Sustainability in e-commerce packaging: a review. *Journal of Cleaner Production, 280,* 124314. https://doi.org/10.1016/j.jclepro.2020.124314

Firmansyah, M.R., Sumarwan, U., & Ali, M.M. (2021). Marketing mix, brand equity, and purchase decision of packaged rice products. *Jurnal Manajemen & Agribisnis, 18*(3), 240-251. https://doi.org/10.17358/jma.18.3.240

Forman, C., Ghose, A., & Goldfarb, A. (2009). Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management Science, 55*(1), 47–57. https://doi.org/10.1287/mnsc.1080.0932

Grosse, E. H., Glock, C. H., & Neumann, W. P. (2017). Human factors in order picking: a content analysis of the literature. International Journal of Production Research, 55(5), 1260–1276. https://doi.org/10.1080/00207543.2016.1186296

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.

Hand, D. J. (2018). Statistical challenges of administrative and transaction data. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 181*(3), 555–605. https://doi.org/10.1111/rssa.12315

Huang, Y., & Bronnenberg, B. J. (2023). Consumer transportation costs and the value of e-commerce: Evidence from the Dutch apparel industry. *Marketing Science, 42*(5), 984–1003. https://doi.org/10.1287/mksc.2022.1421

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30*(2), 199–218. https://doi.org/10.1086/376806

Jewell, C. (2019). *Leveraging Indonesia's creative economy.* WIPO Magazine.

Kondo, F. N., & Okubo, T. (2022). Understanding multi-channel consumer behavior: A comparison between segmentations of multi-channel purchases by product category and overall products. *Journal of Retailing and Consumer Services, 64*, Article 102792. https://doi.org/10.1016/j.jretconser.2021.102792

Kotler P., Armstrong G. 2018. *Principles of Marketing*. 17th ed. United Kingdom: Pearson.

Kumar, V., & Venkatesan, R. (2005). Who are the multichannel shoppers and how do they perform? Correlates of multichannel shopping behavior. *Journal of Interactive Marketing, 19*(2), 44–62. https://doi.org/10.1002/dir.20034

Laudon, K.C., & Traver, C.G. (2021). *E-commerce 2020-2021: Business, Technology, Society* (16). United Kingdom: Pearson.

Martin, J., Nenycz-Thiel, M., Dawes, J., Tanusondjaja, A., Cohen, J., McColl, B., & Trinh, G. (2020). Fundamental basket size patterns and their relation to retail performance. *Journal of Retailing and Consumer Services, 54,* 102032. https://doi.org/10.1016/j.jretconser.2020.102032

Ministry of Creative Economy/Creative Economy Agency of the Republic of Indonesia. (2025, September 16). *Government encourages the utilization of AI to drive the growth of the national creative economy.* https://ekraf.go.id/news-en/government-encourages-the-utilization-of-ai-to-drive-the-growth-of-the-national-creative-economy

Morwitz, V. (2014). Consumers' purchase intentions and their behavior. *Foundations and Trends in Marketing, 7*(3), 181–230. https://doi.org/10.1561/1700000036

Nasution, A.E., & Putri, L.P. (2019). Analisis faktor faktor yang mempengaruhi keputusan konsumen membeli online. *Prosiding FRIMA (Festival Riset Ilmiah Manajemen dan*

*Akuntansi)*, *2*, 491-497. https://prosidingfrima.digitechuniversity.ac.id/index.php/prosidingfrima/article/view/73

Pasharibu, Y., Soerijanto, J.A., & Jie, F. (2020). Intention to buy, interactive marketing, and online purchase decisions. *Jurnal Ekonomi dan Bisnis*, *23*(2), 341-356. https://doi.org/10.24914/jeb.v23i2.3398

Rafiq, M., & Ahmed, P. K. (1995). Using the 7Ps as a generic marketing mix: An exploratory survey of UK and European marketing academics. *Marketing Intelligence & Planning*, *13*(9), 4–15. https://doi.org/10.1108/02634509510097793

Suyanto, A. M. A., & Dewi, D. G. (2023). Marketing mix on purchase intention and its impact on the decision to purchase Somethinc products. *International Journal of Professional Business Review*, *8*(10), e03779. https://doi.org/10.26668/businessreview/2023.v8i10.3779

Widyastuti, A.N., Pujiharto, Tubastuvi, N., & Santoso, S.B. (2020). The effect of marketing mix on purchase decisions. *Jurnal Manajemen Bisnis*, *11*(2), 164-176. https://doi.org/10.18196/mb.11295

Wirtz, J., & Lovelock, C. (2022). *Services marketing: People, technology, strategy* (9th ed.). World Scientific.

Wooldridge, J. M. (2020). *Introductory Econometrics: A Modern Approach* (7th ed.). Cengage Learning.