Journal of
Theoretical and Applied Electronic Commerce Research

MDPI

Article

Latent Class Segmentation of Cross-Border E-Commerce Consumers: Evidence from Korean AliExpress and Temu Users

Minjung Roh

Division of Business Administration, College of Business Administration, Pukyong National University, 45 Yongso-ro, Nam-gu, Busan 48513, Republic of Korea; mroh@pknu.ac.kr

Abstract

Chinese cross-border e-commerce (CBEC) platforms have rapidly reshaped contemporary digital retail, achieving substantial adoption. However, prior research on consumer heterogeneity has yet to sufficiently examine how this fast-growing consumer base is segmented across these platforms' broad category spectrum. The present study accordingly applies latent class analysis to 16 product purchase indicators from 600 Korean AliExpress or Temu users, and conducts bias-adjusted three-step analyses that relate segment membership to demographic and platform covariates and to distal outcomes in shopping behavior and platform awareness channels. Five qualitatively distinct segments emerge: Minimal, Fashion, Household, Auto & Home, and Omnivore shoppers. Omnivore shoppers, the smallest segment, register the highest per-transaction expenditure, satisfaction, and active search. Auto & Home shoppers form a second active-search tier for product search and deal seeking. Household shoppers, the lowest-income tier, show greater receptiveness to offline broadcast advertising. Fashion shoppers have the highest share of consumers aged younger than 20 years and the lowest satisfaction. Minimal shoppers, the largest segment, have the lowest per-transaction expenditure and item count. Platform affiliation crosscuts these segments, with Minimal and Auto & Home shoppers favoring AliExpress and the remaining three Temu. Consumer heterogeneity on CBEC platforms thus forms a clearly differentiated subpopulation structure.

Keywords: cross-border e-commerce; latent class analysis; consumer segmentation; person-centered analysis; consumer heterogeneity

Check for updates

Academic Editors: Yu-Wei Chang, Po-Ya Chang and Jiahe Chen

Received: 10 June 2026
Revised: 23 July 2026
Accepted: 27 July 2026
Published: 3 August 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Korean consumers have rapidly adopted cross-border e-commerce (CBEC) platforms as a meaningful component of their everyday retail behavior [1]. Over the past several years, Chinese-origin cross-border platforms have introduced Korean consumers to a distinctive retail model that offers an unusually broad spectrum of consumer goods categories, ranging from fashion and household items to electronics, automotive accessories, and tools. Their accelerated uptake represents a substantial development in contemporary Korean digital retail.

The rise of these platforms has had substantive consequences for marketers, retailers, and the platforms [2]. Understanding how consumers use cross-border platforms—what they purchase, how much they spend, how they become aware of the platform, and how their behavior varies across consumer types—is necessary for platform marketers to make strategic decisions about category investment and promotional structure. CBEC platforms operate within a platform-mediated retail context in which consumers, brands, and the

platform shape value creation through ongoing interactions [3]. However, despite the strategic importance of understanding consumer-side variations on these platforms, the empirical literature on CBEC has primarily focused on advancing the understanding of adoption drivers and purchase antecedents [4,5], leaving how such heterogeneity organizes into identifiable consumer segments across the broad category spectrum of contemporary platforms insufficiently examined.

Three related gaps emerge from this analytical orientation. First, the methodological tradition dominating CBEC consumer research relies on variable-centered approaches that assume a single underlying population and, thus, cannot identify qualitatively distinct consumer subpopulations [4,5]. Second, prior segmentation work on e-commerce consumer populations has proceeded through attitudinal or psychographic measures rather than behavioral indicators [6,7]; the only precedent for segmenting based on cross-category purchase behavior employed a narrower category set than that provided by contemporary cross-border platforms [8]. Third, although consumer information search has been theoretically structured for decades [9] and has recently been confirmed to vary substantially across consumers on Chinese e-commerce platforms [10], whether active-search-related awareness channels provide a meaningful basis for differentiating segments on CBEC platforms remains unexplored.

This study addresses these gaps through a person-centered empirical examination of consumer heterogeneity on CBEC platforms in the Korean market. Drawing on a sample of 600 Korean users of AliExpress or Temu, the analysis applies a bias-adjusted latent class analysis (LCA) to 16 product purchase indicators spanning a broad spectrum of consumer-good categories, capturing the breadth of contemporary CBEC platforms. The analytical approach proceeds through three integrated stages—class enumeration, covariate profiling, and distal-outcome comparison (Section 3.3). The combined analytical approach yields integrated profiles of each consumer segment that link purchase-portfolio identity to demographic, behavioral, and channel reception characteristics, thereby rendering the heterogeneity of this consumer base visible in a directly interpretable, segment-level form.

This study contributes to the literature on interactive marketing and CBEC in four ways. First, it extends person-centered segmentation research into the CBEC context by characterizing how this consumer base organizes into qualitatively distinct subpopulations. Second, it establishes cross-category purchase behavior as a substantively informative segmentation basis—one that captures portfolio-level consumer heterogeneity beyond what attitudinal and perceptual segmentation bases achieve. Third, it shows that active search heterogeneity differentiates consumer segments along a dimension in which qualitatively distinct types are systematically positioned. Finally, the study provides segment-specific practical implications for cross-border platforms, grounded in the integrated profiles yielded by the three-stage analysis.

The remainder of this paper is organized as follows. Section 2 reviews the literature on interactive marketing and consumer heterogeneity, variable- and person-centered approaches in CBEC consumer research, segment-level analysis of e-commerce platforms, and consumer information search as a basis for differentiation, and articulates three research questions. Section 3 describes the data, measures, and analytical procedures used in this study. Section 4 presents the empirical findings from the three analytical stages. Section 5 discusses the study's theoretical and practical implications, considers its limitations, and suggests directions for future research. Finally, Section 6 concludes the paper.

https://doi.org/10.3390/jtaer21080249

2. Literature Review

2.1. Interactive Marketing and Consumer Heterogeneity

Interactive marketing represents a shift from broadcast communication toward bilateral individual-level exchanges between firms and consumers [11]. It relies on three communicative capabilities: addressing an individual, recording the individual's response, and responding to the individual based on the information obtained. The underlying orientation is inherently person-centered, in that the paradigm presupposes systematic individual differences among consumers, which are presumed to be identifiable and actionable by the firm.

Despite this person-centered orientation embedded in its foundational definitions, the empirical literature has largely treated platform users as a homogeneous population whose responses to interactive marketing instruments can be adequately characterized by average effects. The body of work has considerably expanded in both scale and scope; however, its empirical findings resist reconciliation across studies and settings [12]. Such inconsistency indicates heterogeneity: when aggregate-level analyses yield divergent conclusions across samples, the underlying populations plausibly differ in ways that aggregate framing cannot reveal.

A small but consequential line of work within the interactive marketing tradition has explicitly addressed this gap. Marketing-mix elasticities substantially vary across latent-class-derived segments of digital music consumers, with conventional segmentation procedures—median splits and finer quantile divisions—proving unable to isolate the highest-revenue subgroups identified by the latent-class model [13]. The same analysis further documents that heavy digital music users are the least sensitive to price and most sensitive to advertising, a pattern that contradicts the empirical generalizations developed in traditional consumer goods contexts. The broader mixture-modeling tradition, which encompasses LCA and related person-centered techniques such as latent growth curve modeling, has continued to uncover meaningful subpopulation structures across consumer contexts, including consumer packaged goods adoption [14], establishing the heterogeneity-segmentation orientation as an active research stream rather than an isolated finding.

CBEC platforms operating in the Korean market constitute an appropriate empirical setting for extending this line of inquiry; however, the structure of consumer heterogeneity within these platforms has received little systematic examination. Addressing this question requires methodological tools designed to capture the subpopulation structure rather than tools that estimate average effects across an assumed homogeneous user base, which is reviewed in the next subsection.

2.2. Variable-Centered and Person-Centered Approaches in Cross-Border E-Commerce Consumer Segmentation

Consumer segmentation research can be approached through two methodologically distinct traditions: conventionally labeled variable-centered and person-centered [15]. The variable-centered tradition examines how a set of variables relates to one another within a population assumed to be homogeneous in those relations, yielding a single set of parameters intended to characterize the sample as a whole; regression and structural equation modeling are characteristic of this tradition. Conversely, the person-centered tradition proceeds from the converse assumption that the population comprises latent subgroups whose internal patterns of association qualitatively differ and seeks to identify those subgroups directly from the data; LCA, latent profile analysis, and cluster-analytic methods are characteristic of this tradition [16].

Variable-centered approaches such as regression and structural equation modeling have been the dominant analytical tools in CBEC consumer research [4,5]. Their guiding

premise is that the sample under analysis is derived from a unitary population whose behavioral regularities can be summarized through a common set of estimated parameters [17]. Accordingly, inferential attention is directed at the average associations among variables, with the resulting estimates used to characterize the population in the aggregate. This framing is well calibrated to populations that are genuinely homogeneous, but imposes a structural limitation when consumers cluster into qualitatively distinct behavioral types. Under such circumstances, subgroup-level patterns are absorbed into aggregate means rather than being identified as distinct entities, and the heterogeneity that segmentation research aims to characterize remains analytically invisible.

Person-centered approaches abandon the homogeneity assumption and instead model the population as a composite of unobserved subgroups, each governed by its own parameter set [17]. Within this family, LCA offers a model-based foundation; it specifies an explicit probabilistic generating mechanism—a finite mixture of class-conditional distributions from which the observed data are assumed to arise—and estimates the resulting model by maximum likelihood, contrasting the distance-minimization heuristics that underpin conventional clustering algorithms [18]. Simulation evidence shows that latent class clustering identifies true subpopulation structures with an accuracy approaching that of discriminant analysis applied with full knowledge of class membership and substantially outperforms heuristic clustering techniques such as K-means [18]. The model-based grounding also confers practical advantages relevant to consumer segmentation: formal information criteria for selecting the number of classes, probabilistic rather than deterministic class assignment, and invariance to the variable-scaling decisions that distort distance-based solutions.

Therefore, CBEC consumer populations warrant analysis as a question of subpopulation structure rather than as average effects. Consumer heterogeneity in interactive marketing settings has been empirically established [13,14]; however, whether such heterogeneity assumes the form of qualitatively distinct consumer types on CBEC platforms and how any such types differ in their purchase patterns and behavioral outcomes are questions that variable-centered analyses are structurally unequipped to address. LCA, with its capacity to capture the subpopulation structure directly from observed data, is the appropriate methodological response to questions of this form.

2.3. Segment-Level Analysis of E-Commerce Consumer Populations

Prior research comprises person-centered segmentation analyses of consumer populations in e-commerce contexts, which have accumulated since the early 2000s [8]. An early application to online retailing identified a three-segment structure from web shoppers' cross-category purchase patterns and then profiled the resulting segments along two dimensions: demographics and benefits sought. Demographic attributes were largely undiagnostic, whereas benefit-related perceptions discriminated sharply among segments and revealed differentiated patterns of attribute importance and web-performance evaluation [8]. This line of work demonstrates that consumer heterogeneity in online purchasing is both structured and analytically tractable, and that segmentation analysis reveals patterns that aggregate descriptions of online shoppers cannot reproduce.

Subsequent applications extended the segment-level analysis to multichannel retail. Among 364 Dutch consumers, LCA identified three multichannel shopper segments—multichannel enthusiasts, uninvolved shoppers, and store-focused consumers—whose membership was predicted by psychological, economic, and sociodemographic covariates, with the broad segment structure proving stable across product categories even though the strength of specific covariate effects varied [7]. A study of 1325 consumers in Mainland China extended the same analytical orientation to multichannel, multistage consumption, identifying innovative and conventional consumer segments whose valuations of online

versus offline channels at the information search and purchase stages differed systematically across both search and experience product categories [19]. A latent class regression analysis of Korean over-the-top media users identified three behaviorally distinct segments from usage pattern indicators and used demographic covariates to predict segment membership [20]. The methodological positioning across these studies treats latent class modeling as an analytical response to the recognition that consumers exposed to nominally similar e-commerce environments respond in qualitatively distinct ways.

Segment-level analysis has also been applied to specific e-commerce platform contexts. Among 1027 Turkish university students, the psychographic segmentation of business-to-consumer e-commerce users yielded four consumer types—shopping lovers, direct purchasers, suspicious browsers, and incompetent consumers—whose subsequent shopping behaviors differed across segments [6]. This study's methodological contribution lies in its demonstration that aggregate-level findings on e-commerce consumer behavior can systematically misrepresent the behavioral patterns of identifiable consumer subgroups, positioning segment-level analysis as a corrective to single-population assumptions. A study of 309 Facebook clothing shoppers at a large Asian university used mixture modeling to identify three consumer segments: social patrons, wary explorers, and sporadic explorers [21]. This latent class approach was motivated by the fact that a priori segmentation based on demographic or psychographic variables does not address the unobserved heterogeneity revealed by mixture modeling.

Recent clickstream evidence from JD.com, a major Chinese e-commerce platform, further documents that consumer-level heterogeneity in search behavior, specifically in the breadth and depth with which consumers evaluate products, is empirically substantial and consequential for platform-side strategies [10], confirming that consumer heterogeneity remains an active empirical concern in contemporary e-commerce settings. Despite this research across various e-commerce platform types, few attempts have specifically applied segment-level LCA to consumer populations on CBEC platforms. Across an even broader e-commerce segmentation literature, the prevailing segmentation bases have been attitudinal or perceptual constructs such as channel attitudes [7], channel value perceptions [19], psychographic factors [6], and UTAUT2 constructs augmented with trust [21]. Only the early-web-era study by Bhatnagar and Ghose [8] employed cross-category purchase behavior as the segmentation basis, with a category set that reflected the technology-heavy online retail landscape of that period rather than the broad consumer goods spectrum supplied by contemporary CBEC platforms. The methodological apparatus available at the time also predated the bias-adjusted three-step procedures [22] that contemporary latent class practices now incorporate for covariate and distal outcome analyses.

2.4. Consumer Information Search as a Basis for Differentiation

Consumer information search is a behavior that systematically varies in magnitude across consumers as a function of factors operating at the individual level. Foundational work on automobile purchasing distinguished two dimensions of consumer knowledge with opposing effects on external search: usable prior information, which reduces the need for external search, and prior memory structure, which prompts a more extensive search. The cost of external search has emerged as a further suppressor of such search activities [23]. The core insight is that search effort is neither random nor uniform across consumers but systematically structured by consumer-level differences in prior knowledge and the cost of external search.

This construct extends beyond impending purchase contexts. Consumers engage in ongoing searches disconnected from any immediate purchase decision, driven by hedonic and recreational motives rather than by direct informational ones [24]. Enduring product

involvement often serves as the central antecedent of such ongoing searches and consumers who search heavily act as influential transmitters of product information in the marketplace. These contributions characterize search effort as a behavioral construct with structured antecedents and distinct behavioral expressions across purchase and non-purchase contexts.

Schmidt and Spreng's [9] integrative model consolidates this literature into a parsimonious framework. The model organizes the determinants of external information search through four mediating constructs—ability, motivation, costs, and benefits—and comprehensively incorporates nearly 20 determinants, while remaining parsimonious through these four organizing factors. The substantive implication is that consumer-side variations in search effort reflect underlying differences in how individuals are positioned regarding ability, motivation, costs, and benefits. This framework supports treating active search as a unified behavioral dimension whose variation across consumers is theoretically grounded rather than analytically arbitrary.

Recent empirical work confirms that prepurchase consumer search remains a central concern in contemporary consumer research, with observed search patterns implying substantial heterogeneity across consumers. A review of pre-purchase search data documents two regularities across product categories [25]: search incidences substantially exceed purchase incidences, with apparel browsing data showing a ratio approaching tenfold [26], and average consideration sets remain remarkably small, with reported sizes typically below three [27–29].

Beyond aggregate empirical regularities, platform-level evidence further establishes that consumer search heterogeneity is a substantive feature of e-commerce. Clickstream analysis of users on JD.com, a major Chinese e-commerce platform, distinguishes naïve from sophisticated consumers based on observed differences in search breadth and depth, with sophisticated consumers examining a larger set of products and inspecting each product more thoroughly than their naïve counterparts [10]. The same study demonstrates that a platform-side search design—how products are presented, ordered, and recommended to users—must accommodate this consumer-side heterogeneity to be effective, with the platform's optimal search-cost setting contingent on the relative composition of these two consumer types within its user base. The implication is that consumer search behavior on e-commerce platforms is substantively heterogeneous and directly affects the design choices through which platforms manage that heterogeneity.

Together, these studies establish active consumer search as a structured behavioral construct. The magnitude of search varies across consumers with prior knowledge, search costs, and expected benefits; the construct itself spans both purchase-related and ongoing search activity; and the resulting variation across consumers is systematic rather than incidental. Whether such heterogeneity is sufficiently pronounced to differentiate consumer segments within the CBEC setting—and, if so, how such segments covary with platform-level behavioral outcomes—is the empirical question this study directly addresses.

2.5. Research Gap and Research Questions

Four gaps emerge from the preceding review. First, although the contemporary interactive marketing paradigm rests on a person-centered foundation [11], the empirical literature in this tradition has predominantly examined platform users as a homogeneous population whose responses can be summarized through average effects, leaving consumer heterogeneity on CBEC platforms substantially under-examined as a structural feature of the user base.

Second, the methodological tradition that dominates CBEC consumer research relies on variable-centered approaches, most notably structural equation modeling [4,5], that assume that all individuals are drawn from a single underlying population for which a

single set of averaged parameters can be estimated [17]. Therefore, it cannot identify the qualitatively distinct consumer subpopulations that segmentation analysis is designed to reveal.

Third, although prior segmentation work has applied LCA to consumer populations on various e-commerce platform types, few attempts have specifically applied this analytical orientation to consumer populations on CBEC platforms; the only precedent that has segmented consumers on cross-category purchase behavior [8] pre-dates the contemporary cross-border setting in two important respects: its narrower category set reflects an earlier online retailing landscape, and its analytical apparatus predates the methodological refinements [22] that latent class practices now incorporate.

Fourth, although consumer information search has been theoretically structured for decades [9] and has been recently confirmed to substantially vary across consumers on Chinese e-commerce platforms [10], whether active search behavior provides a meaningful basis for differentiating consumer segments on CBEC platforms remains empirically unexamined.

Together, these gaps motivate a person-centered empirical examination of consumer heterogeneity on CBEC platforms, which connects the contemporary interactive marketing paradigm to the structural realities of cross-border platform usage. Such an examination must establish whether CBEC users constitute the homogeneous population that aggregate analyses implicitly assume, or whether qualitatively distinct consumer types coexist within the same platform context and, if so, how those types are predicted by observable consumer characteristics and differ in the behavioral and awareness-channel measures that are most consequential for interactive marketing strategies. Addressing these questions requires three integrated analytical components: identifying latent consumer segments from observed product purchase patterns; examining how demographic characteristics and platform affiliation predict segment membership; and characterizing how segments differ in shopping behavior and platform awareness channels.

Accordingly, this study addresses the following research questions:

RQ1. What distinct consumer segments emerge from product purchase patterns on CBEC platforms?

RQ2. How do demographic characteristics and platform affiliation predict segment membership?

RQ3. How do consumer segments differ in shopping behavior and platform awareness channels?

3. Materials and Methods

3.1. Dataset and Sampling Description

The empirical basis for this investigation is a de-identified secondary dataset licensed from Opensurvey (https://www.opensurvey.co.kr, accessed on 3 October 2024), a Korean online research firm that maintains a registered consumer panel and operates under ISMS-P certification for information security and privacy management. Eligibility for the present study was restricted to panelists who had used a Chinese CBEC platform, AliExpress or Temu, at least once within the 30 days prior to fieldwork, ensuring that all respondents had recent first-hand experience with the platforms under study. Fieldwork was conducted via a mobile-optimized questionnaire over a two-day window (9–10 April 2024). The full Opensurvey 2024 release comprised 1200 panelists sampled across both domestic and CBEC contexts; only the 600 respondents who reported recent use of AliExpress or Temu were routed to the CBEC-specific question block and constituted the analytic sample for the present study, with equal representation of the two platforms. The demographic composition of the participants is summarized in Table 1.

Table 1. Sample characteristics.

Variable	AliExpress (n = 300)	Temu (n = 300)	Overall (n = 600)
Gender			
Male	198 (66.0%)	176 (58.7%)	374 (62.3%)
Female	102 (34.0%)	124 (41.3%)	226 (37.7%)
Age			
10s	19 (6.3%)	10 (3.3%)	29 (4.8%)
20s	49 (16.3%)	51 (17.0%)	100 (16.7%)
30s	64 (21.3%)	58 (19.3%)	122 (20.3%)
40s	87 (29.0%)	57 (19.0%)	144 (24.0%)
50s	52 (17.3%)	68 (22.7%)	120 (20.0%)
60s	29 (9.7%)	56 (18.7%)	85 (14.2%)
Income (monthly household, KRW)			
1 (<2 million)	17 (5.7%)	15 (5.0%)	32 (5.3%)
2 (2 to <3 million)	21 (7.0%)	31 (10.3%)	52 (8.7%)
3 (3 to <4 million)	40 (13.3%)	26 (8.7%)	66 (11.0%)
4 (4 to <5 million)	45 (15.0%)	45 (15.0%)	90 (15.0%)
5 (5 to <6 million)	50 (16.7%)	49 (16.3%)	99 (16.5%)
6 (6 to <7 million)	37 (12.3%)	27 (9.0%)	64 (10.7%)
7 (7 to <8 million)	24 (8.0%)	31 (10.3%)	55 (9.2%)
8 (8 to <9 million)	19 (6.3%)	19 (6.3%)	38 (6.3%)
9 (9 to <10 million)	15 (5.0%)	19 (6.3%)	34 (5.7%)
10 (≥10 million)	32 (10.7%)	38 (12.7%)	70 (11.7%)
Marital status			
Unmarried	114 (38.0%)	102 (34.0%)	216 (36.0%)
Married	175 (58.3%)	190 (63.3%)	365 (60.8%)
Other (divorced/widowed)	11 (3.7%)	8 (2.7%)	19 (3.2%)

3.2. Measures

3.2.1. LCA Indicators

Latent class indicators measured product purchase experience on CBEC platforms. Respondents indicated whether they had ever purchased products in the 17 categories on the platform (0 = not purchased, 1 = purchased): fashion clothing, fashion accessories, cosmetics/perfume, sports/leisure goods, household essentials, kitchen supplies, maternity/childcare, furniture/interior, electronics/digital, food/groceries, automotive accessories, tools/hardware, pet supplies, health/medical supplies, stationery/hobby goods, books/media, and an "other" category with an optional free-text specification. The 17th "other" category was excluded from analysis due to the heterogeneous nature of free-text responses, yielding 16 binary indicators for the LCA.

3.2.2. Covariates

Five variables served as covariates in the R3STEP analysis to predict latent class membership. Gender was coded as a binary variable (0 = male, 1 = female). Age was measured in decade intervals (10s, 20s, 30s, 40s, 50s, and 60s). Marital status was categorized as unmarried, married, or divorced/widowed. Income level was measured as monthly household income on a ten-point scale: 1 (<KRW 2 million), 2 (KRW 2 to <3 million), 3 (KRW 3 to <4 million), 4 (KRW 4 to <5 million), 5 (KRW 5 to <6 million), 6 (KRW 6 to <7 million), 7 (KRW 7 to <8 million), 8 (KRW 8 to <9 million), 9 (KRW 9 to <10 million), and 10 (≥KRW 10 million). At the 10 April 2024 exchange rate of 1363.66 KRW per USD, KRW 2 million corresponded to approximately US$1467 and KRW 10 million to approximately US$7333. Platform was coded as a binary variable, indicating whether the respondent was an AliExpress (0) or Temu (1) user.

3.2.3. Distal Outcomes

Two sets of distal outcome variables were examined using the Bolck–Croon–Hagenaars (BCH) method. The first set comprises three shopping behavior variables. Per-transaction expenditure was measured on an eight-point scale: 1 (<KRW 30,000), 2 (KRW 30,000 to <50,000), 3 (KRW 50,000 to <100,000), 4 (KRW 100,000 to <200,000), 5 (KRW 200,000 to <300,000), 6 (KRW 300,000 to <400,000), 7 (KRW 400,000 to <500,000), and 8 (≥KRW 500,000). At the 10 April 2024 exchange rate of 1363.66 KRW per USD, KRW 30,000 corresponded to approximately US$22 and KRW 500,000 to approximately US$367. Items per transaction were measured on a six-point scale: 1 (1–2 items), 2 (3–4 items), 3 (5–6 items), 4 (7–8 items), 5 (9–10 items), and 6 (more than 10 items). Overall satisfaction was measured on a five-point scale (1 = very dissatisfied, 5 = very satisfied). Although these variables were ordinal in nature, they were treated as continuous in the primary BCH analysis, with the ordinal treatment retained as a robustness check.

The second set comprised six binary platform awareness channel variables. Respondents indicated all the channels through which they first became aware of the platform from seven options (select all that apply): product search on the app/web, searching for deals/benefits, TV/radio/offline ads, SNS/social media/YouTube ads, SNS/social media/YouTube reviews, acquaintance recommendations, and "other" (free-text specification). The seventh "other" option was excluded from analysis due to the heterogeneous nature of free-text responses, yielding six binary variables (0 = not selected, 1 = selected).

3.3. Analytical Approach

The analysis was conducted in three sequential stages (Figure 1): class enumeration, covariate profiling, and distal outcome comparison.

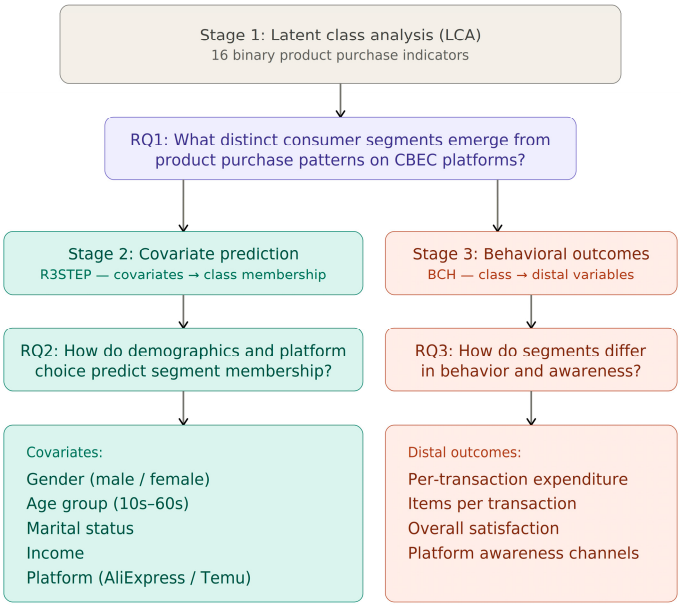


Figure 1. Three-stage analytical framework for consumer segmentation in cross-border e-commerce.

In the first stage, a series of latent class models, ranging from one to ten classes, were fitted to 16 binary product purchase indicators. The number of classes was determined by evaluating multiple fit criteria jointly, as each criterion addresses a distinct dimension of model adequacy. The Bayesian Information Criterion (BIC) balances fit against parsimony and serves as the primary basis for comparing competing enumerations. As information criteria do not formally test whether an additional class yields a statistically significant improvement, the bootstrap likelihood ratio test (BLRT) provides this test. Entropy R^2 and

classification error capture a third dimension—the precision with which individuals can be assigned to classes—which bears directly on the present design, since both the covariate profiling of the second stage (R3STEP) and the distal outcome comparisons of the third stage (BCH) correct for classification error, and the reliability of these corrections deteriorates as class separation weakens [17]. Finally, bivariate residual (BVR) diagnostics evaluate local, rather than global, adequacy. Latent class models rely on the local independence assumption: within any given class, the indicators are uncorrelated, so that the class structure itself accounts for all observed associations among them. The conditional probabilities that constitute class profiles are interpretable only insofar as this assumption holds; residual within-class association between an indicator pair signals that part of the apparent profile structure reflects indicator-level overlap rather than genuine class differences. Because the BIC, BLRT, and entropy summarize fit in aggregate across the full response pattern, a model can perform well on all of these criteria while leaving substantial residual dependence in specific indicator pairs—a possibility borne out in the present enumeration, where the models favored by the BIC alone failed the local independence check (Section 4.1). BVR diagnostics, therefore, supply the local verification that the global criteria cannot. The final model was required to meet all of the following criteria: a significant BLRT, an entropy R^2 of at least 0.600, a classification error below 0.200, and all BVR values below 3.840. Among the models satisfying these requirements, the BIC was used to identify the most parsimonious solution. Residual indicator-pair associations exceeding the BVR threshold were addressed by freeing direct effects between the affected pairs, rather than by increasing the number of classes. This choice reflects the source of the dependence: residual association confined to specific indicator pairs—such as the complementary purchases of fashion clothing and fashion accessories—constitutes an indicator-level phenomenon rather than evidence of an additional consumer type. Adding classes to absorb such an association, therefore, risks overextraction—the retention of spurious classes that mirror item-pair associations rather than genuine consumer types. Freeing the direct effect instead models the association explicitly while leaving the class structure intact [30].

In the second stage, demographic and platform covariates (gender, age, marital status, income, and platform) were linked to class membership using the R3STEP approach [22]. R3STEP estimates a bias-adjusted multinomial logistic regression that accounts for classification errors while leaving the latent class measurement model undisturbed, thereby preventing the distortion of the class structure that can occur when covariates are directly incorporated into the measurement model.

In the third stage, shopping behavior variables (per-transaction expenditure, items per transaction, and overall satisfaction) and the six binary platform awareness channel indicators were compared across the identified classes using the BCH procedure [31,32]. BCH employs inverse-classification error weights to correct for the uncertainty inherent in modal class assignment, yielding unbiased class-specific outcome estimates. To assess sensitivity to measurement scale assumptions, shopping behavior analyses were replicated using an ordinal specification. All models were estimated using Latent GOLD 6.1 (Statistical Innovations, Arlington, MA, USA) [30].

Generative AI assistance was limited to language editing and reference formatting; all data preparation, modeling, analysis, and interpretation were conducted solely by the author.

4. Results

4.1. Model Selection and Class Identification

To determine the optimal number of latent classes, models ranging from one to ten classes were estimated without direct effects (Appendix A). Although the BIC favored the

four-class solution (8942.471), bivariate residual diagnostics revealed that the three-, four-, and five-class models all violated the local independence assumption, with maximum BVR values of 40.712, 11.438, and 5.410, respectively, each exceeding the critical value of 3.840. Among these, the five-class solution produced the lowest residual dependence (max BVR = 5.410, still exceeding this threshold) and was further supported by a significant BLRT ($p < 0.001$); therefore, it was selected as the basis for direct-effects specification.

Inspection of the no-direct-effects five-class model identified five indicator pairs whose bivariate residuals exceeded 3.840: fashion clothing × fashion accessories (5.007), cosmetics/perfume × pet supplies (4.684), furniture/interior × stationery/hobby goods (5.410), tools/hardware × books/media (5.082), and pet supplies × books/media (4.131). The first pairing reflects within-domain complementarity among fashion items; the remaining four reflect cross-category co-purchase tendencies that are not absorbed by the latent class structure (Appendix B). Once these five pairs were freed as direct effects and all models were re-estimated (Table 2), the maximum BVR for the five-class solution fell to 3.585, confirming that the local independence assumption was satisfied (Appendix C).

Table 2. Model comparison for latent class enumeration (1- to 10-class models with five direct effects).

Classes	LL	BIC	AIC	AIC3	Entropy R^2	Class. Error	Max BVR	Bootstrap p	−2LL Diff	BLRT p	Smallest Class (%)
1	−4735.911	9606.158	9513.822	9534.822	1.000	0.000	188.311	—	—	—	100.000
2	−4393.880	9030.842	8863.759	8901.759	0.899	0.011	49.948	<0.001	684.063	<0.001	7.910
3	−4285.770	8923.371	8681.540	8736.540	0.746	0.067	20.848	0.088	216.219	<0.001	3.830
4	−4227.973	8916.525	8599.947	8671.947	0.670	0.158	10.388	0.100	115.594	<0.001	3.730
5	**−4190.646**	**8950.618**	**8559.291**	**8648.291**	**0.676**	**0.182**	**3.585**	**0.192**	**74.655**	**<0.001**	**3.720**
6	−4167.873	9013.821	8547.747	8653.747	0.699	0.183	4.242	0.158	45.545	0.014	3.690
7	−4150.137	9087.096	8546.274	8669.274	0.689	0.194	4.303	0.166	35.473	0.130	1.740
8	−4134.582	9164.733	8549.163	8689.163	0.715	0.201	3.097	0.162	31.110	0.166	1.910
9	−4121.685	9247.688	8557.370	8714.370	0.718	0.206	2.091	0.084	25.793	0.300	1.070
10	−4105.905	9324.876	8559.810	8733.810	0.731	0.198	1.870	0.030	31.560	0.104	0.520

Note. Five direct effects freed: Fashion clothing × Fashion accessories, Cosmetics/perfume × Pet supplies, Furniture/interior × Stationery/hobby goods, Tools/hardware × Books/media, Pet supplies × Books/media. Bold font indicates the selected model. Class. error = Classification error. BLRT = Bootstrap likelihood-ratio test (500 replications).

Within the direct-effects specification, the five-class model emerged as the optimal solution. The three- and four-class models returned lower BIC values (8923.371 and 8916.525 vs. 8950.618), yet neither met the local independence requirement (maximum BVR = 20.848 and 10.388, respectively); the five-class model was the first solution in which all bivariate residuals fell below the 3.840 threshold (maximum BVR = 3.585). Sequential BLRTs were significant through the five-class transition (all $p < 0.001$), and the entropy R^2 of 0.676 surpassed the 0.60 benchmark for adequate classification quality. Adding a sixth class lowered the AIC (8547.747), but inflated the BIC to 9013.821. Therefore, the five-class model, with a classification error of 0.182, was adopted as the final solution.

4.2. Class Profiles and Description

Table 3 reports the likelihood of purchasing each of the 16 product categories across the five latent classes. Figure 2 presents these profiles in a radar chart. Each class displayed a distinct combination of product endorsement patterns, as described below.

Minimal shoppers (C1; 36.9%) constituted the largest segment and were characterized by selective, functional purchasing concentrated in tools/hardware (0.421), electronics/digital (0.383), automotive accessories (0.367), and stationery/hobby goods (0.334). The purchase probabilities for fashion (clothing: 0.097; accessories: 0.247), cosmetics (0.008), and food/groceries (0.111) were notably low.

Table 3. Conditional probabilities of product purchase by latent class (five-class direct-effects model).

Indicator	C1	C2	C3	C4	C5	Overall
	Minimal Shoppers	**Fashion Shoppers**	**Household Shoppers**	**Auto & Home Shoppers**	**Omnivore Shoppers**	
Fashion clothing	0.097	**0.609**	0.348	**0.571**	**0.995**	0.383
Fashion accessories	0.247	**0.581**	**0.479**	**0.628**	**0.995**	0.463
Cosmetics/perfume	0.008	0.105	0.084	0.175	**0.947**	0.108
Sports/leisure goods	0.243	0.220	0.029	**0.654**	**0.904**	0.292
Household essentials	0.242	0.125	**0.657**	**0.548**	**0.994**	0.360
Kitchen supplies	0.057	0.153	**0.705**	**0.679**	**0.994**	0.328
Maternity/childcare	0.000	0.044	0.059	0.073	**0.896**	0.067
Furniture/interior	0.140	0.131	0.370	**0.620**	**0.994**	0.287
Electronics/digital	0.383	0.025	0.105	**0.455**	**0.994**	0.277
Food/groceries	0.111	0.021	0.164	0.157	**0.984**	0.137
Automotive accessories	0.367	0.013	0.076	**0.827**	**0.816**	0.317
Tools/hardware	**0.421**	0.066	0.195	**0.760**	**0.994**	0.367
Pet supplies	0.032	0.012	0.158	0.072	**0.762**	0.082
Health/medical supplies	0.007	0.050	0.035	0.082	**0.813**	0.065
Stationery/hobby goods	0.334	0.220	**0.468**	**0.570**	**0.995**	0.390
Books/media	0.012	0.000	0.000	0.137	**0.851**	0.058
Class size	**0.369**	**0.260**	**0.171**	**0.163**	**0.037**	

Note. Values are conditional probabilities of product purchase given class membership, estimated using the five-class model with five direct effects. Class size is the estimated proportion of the sample in each class. The bold font indicates the higher purchase probabilities that characterize each class.

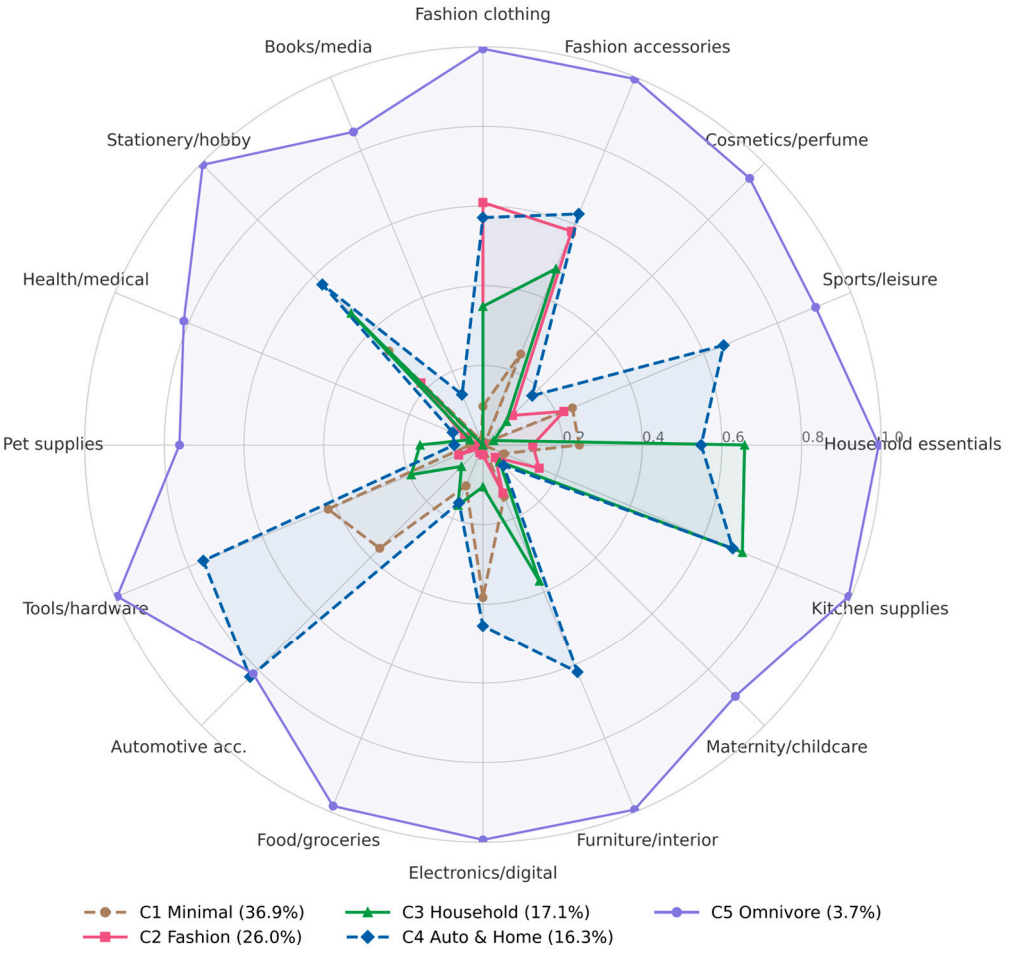


Figure 2. Conditional purchase probabilities across the 16 product categories for the five latent classes. (Values are conditional probabilities ranging from 0 to 1; each axis represents one product category.).

Fashion shoppers (C2; 26.0%) displayed a consumption profile anchored in appearance-related categories. Aside from the small Omnivore segment (C5), which exhibited uniformly

elevated probabilities across nearly all categories, C2 recorded the highest probabilities for fashion clothing (0.609) and fashion accessories (0.581). All non-fashion categories recorded probabilities below 0.250, and the near-zero probabilities for automotive accessories (0.013), electronics/digital (0.025), and food/groceries (0.021) indicated that this segment's cross-border purchasing was narrowly concentrated in fashion.

Household shoppers (C3; 17.1%) were defined by a high endorsement of household-oriented categories such as kitchen supplies (0.705), household essentials (0.657), stationery/hobbies (0.468), and furniture/interior (0.370). Fashion accessories (0.479) and fashion clothing (0.348) were also prominent, distinguishing this class from the purely utilitarian C1 profile. Automotive accessories (0.076), electronics/digital (0.105), and tools/hardware (0.195) were comparatively low.

Auto & Home shoppers (C4; 16.3%) exhibited the broadest product portfolio after C5, with elevated probabilities across automotive accessories (0.827), tools/hardware (0.760), kitchen supplies (0.679), sports/leisure goods (0.654), furniture/interior (0.620), fashion accessories (0.628), and fashion clothing (0.571). While this class purchased broadly across both functional and appearance-related product domains, the probabilities remained low for pet supplies (0.072), health/medical supplies (0.082), maternity/childcare (0.073), and books/media (0.137), indicating a consumption profile anchored in automotive, home improvement, and lifestyle goods rather than in specialty categories.

Omnivore shoppers (C5; 3.7%) represented a small but distinctive segment, with conditional probabilities exceeding 0.800 in 15 of the 16 categories. Eleven of the 16 categories exceeded 0.900, and even categories with low base rates in the overall sample—such as maternity/childcare (0.896), automotive accessories (0.816), pet supplies (0.762), health/medical supplies (0.813), and books/media (0.851)—recorded high probabilities in this class.

Viewed comparatively, the five profiles thus differ not only in how much of the category spectrum they cover—ascending from the concentrated C2 to the near-universal C5—but also in which portion of it they occupy, with C1 and C4 anchored in functional domains, C2 in appearance-related ones, and C3 in household ones.

4.3. Covariates of Class Membership

Table 4 reports the R3STEP multinomial logistic regression results for the effects of demographic characteristics and platform affiliation on latent class membership (L^2 = 476.692, df = 556, p = 0.990; Entropy R^2 = 0.231). All five covariates reached statistical significance.

Gender was strongly associated with class membership (*Wald* = 38.095, $p < 0.001$). Minimal shoppers (C1) were overwhelmingly male (92.34%), as were Auto & Home shoppers (C4; 82.45%). Household shoppers (C3) were predominantly female (76.09%), as were Fashion shoppers (C2; 62.51%). Omnivore Shoppers (C5) showed a relatively balanced distribution (54.21% female).

Age was the strongest predictor (*Wald* = 57.995, $p < 0.001$). Fashion shoppers (C2) had the highest concentration of teenagers (10.16%), while Auto & Home shoppers (C4) were concentrated in the 40s age group (37.64%) with virtually no teenage representation (0.02%). Minimal (C1) and Household (C3) shoppers showed relatively even age distributions across the 20s–60s range. Omnivore shoppers (C5) had the highest concentration in their 30s (31.71%).

Table 4. R3STEP multinomial logistic regression: covariate effects on class membership.

Covariate	C1	C2	C3	C4	C5	Wald	*p*
	Minimal Shoppers	**Fashion Shoppers**	**Household Shoppers**	**Auto & Home Shoppers**	**Omnivore Shoppers**		
Panel A: Coefficients							
Gender [a]						38.095	<0.001
Male	1.164	−0.540	−0.967	0.694	−0.352		
Female	−1.164	0.540	0.967	−0.694	0.352		
Age [b]						57.995	<0.001
10s	0.608	1.396	−0.693	−2.698	1.387		
20s	−0.239	−0.372	−0.367	0.953	0.025		
30s	−0.427	−0.298	0.075	0.458	0.192		
40s	0.005	−0.501	0.344	0.847	−0.696		
50s	−0.244	−0.059	0.272	0.533	−0.501		
60s	0.296	−0.166	0.369	−0.092	−0.406		
Marital status [c]						38.012	<0.001
Unmarried	0.634	−0.115	−0.341	−0.854	0.675		
Married	0.200	−0.363	−0.705	−0.085	0.953		
Divorced/widowed	−0.834	0.477	1.045	0.939	−1.628		
Income [d]						9.681	0.046
Continuous	0.031	−0.083	−0.183	0.127	0.109		
Platform [e]						25.920	<0.001
AliExpress	0.661	−0.216	−0.354	0.722	−0.813		
Temu	−0.661	0.216	0.354	−0.722	0.813		
Panel B: Class profile (%)						**Overall**	
Gender							
Male	92.340	37.490	23.910	82.450	45.790	62.330	
Female	7.660	62.510	76.090	17.550	54.210	37.670	
Age							
10s	4.870	10.160	1.020	0.020	4.500	4.830	
20s	16.790	19.340	16.290	11.950	18.220	16.670	
30s	17.540	21.840	23.550	17.930	31.710	20.340	
40s	26.770	14.600	22.550	37.640	13.300	24.000	
50s	17.450	20.890	19.740	24.780	18.790	20.000	
60s	16.580	13.170	16.850	7.680	13.490	14.170	
Marital status							
Unmarried	38.780	43.810	35.850	17.610	32.530	36.000	
Married	60.400	53.760	56.970	76.360	67.470	60.840	
Divorced/widowed	0.820	2.430	7.180	6.030	0.000	3.170	
Income							
Mean	5.480	5.230	4.680	6.270	6.700	5.450	
Platform							
AliExpress	63.670	36.850	32.620	68.890	13.200	49.990	
Temu	36.330	63.150	67.380	31.110	86.800	50.010	

Note. Panel A reports the effect-coded multinomial logistic regression coefficients. Wald statistics and *p*-values test the overall effect of each covariate on class membership. Panel B reports the class-specific profile percentages, with the overall sample distribution in the final column. [a] Gender, effect-coded (male, female). [b] Age, effect-coded across six groups (10s, 20s, 30s, 40s, 50s, 60s). [c] Marital status, effect-coded (unmarried, married, divorced/widowed). [d] Income, treated as a continuous variable (1–10 scale). [e] Platform, effect-coded (AliExpress, Temu). Bold font indicates panel headings and column headings.

Marital status (*Wald* = 38.012, $p < 0.001$) showed that unmarried respondents were overrepresented among Fashion shoppers (C2; 43.81%) and underrepresented among Auto & Home shoppers (C4; 17.61%). Conversely, married respondents were concentrated in C4 (76.36%). Divorced or widowed individuals, while comprising only 3.17% of the sample, were disproportionately represented among Household (C3; 7.18%) and Auto & Home (C4; 6.03%) shoppers.

Income also showed a significant but weaker association (*Wald* = 9.681, $p = 0.046$). Auto & Home (C4) and Omnivore (C5) shoppers reported the highest mean incomes (6.270 and 6.700 on the ten-point scale, respectively), while Household shoppers (C3) reported the lowest (4.680).

Platform was significantly associated with class membership (*Wald* = 25.920; $p < 0.001$). AliExpress users were disproportionately represented among Minimal shoppers (C1; 63.67%) and Auto & Home shoppers (C4; 68.89%), while Temu users dominated Fashion

(C2; 63.15%), Household (C3; 67.38%), and Omnivore (C5; 86.80%) shoppers. This asymmetric distribution suggests that the two platforms, despite operating in the same CBEC market, attract qualitatively different consumer segments: AliExpress draws more utilitarian, male-skewed shoppers, while Temu attracts fashion-oriented, household-focused, and omnivorous consumers.

4.4. Shopping Behavior and Platform Awareness Channels

Table 5 reports the BCH distal outcome results for shopping behavior (per-transaction expenditure, items per transaction, and overall satisfaction) and platform awareness channels. All three shopping behavior variables and five of the six awareness channels showed significant class differences.

Table 5. BCH distal outcome analysis: shopping behavior and platform awareness channels.

Outcome	C1	C2	C3	C4	C5	Wald	p	R^2
	Minimal Shoppers	**Fashion Shoppers**	**Household Shoppers**	**Auto & Home Shoppers**	**Omnivore Shoppers**			
Shopping behavior (BCH-adjusted means)								
Per-transaction expenditure	1.571	1.794	1.808	2.131	3.418	41.155	<0.001	0.132
Items per transaction	1.657	1.944	2.649	2.346	2.667	70.280	<0.001	0.171
Overall satisfaction	3.489	3.279	3.458	3.738	3.865	21.759	<0.001	0.038
Platform awareness channels (BCH-adjusted proportions)								
Active search								
Product search on app/web	0.277	0.226	0.410	0.422	0.728	23.549	<0.001	0.055
Searching for deals/benefits	0.136	0.090	0.159	0.269	0.638	34.753	<0.001	0.084
Advertising exposure								
TV/radio/offline ads	0.107	0.111	0.213	0.145	0.631	31.543	<0.001	0.080
SNS/social media/YouTube ads	0.352	0.378	0.380	0.452	0.676	9.135	0.058	0.018
Social information								
SNS/social media/YouTube reviews	0.243	0.286	0.201	0.314	0.593	12.623	0.013	0.027
Acquaintance recommendation	0.270	0.271	0.218	0.343	0.586	11.563	0.021	0.023

Note. Shopping behavior values represent BCH-adjusted class-specific means [31,32]. Platform awareness channel values represent BCH-adjusted class-specific proportions. Bold font indicates outcome domains; italic font indicates groupings within the platform awareness channels.

4.4.1. Shopping Behavior

Per-transaction expenditure differed significantly across classes (*Wald* = 41.155, $p < 0.001$; $R^2 = 0.132$). Omnivore shoppers (C5) spent the most per transaction ($M = 3.418$), followed by Auto & Home shoppers (C4; $M = 2.131$). Minimal shoppers (C1) and Fashion shoppers (C2) spent the least ($M = 1.571$ and 1.794, respectively).

Items per transaction (*Wald* = 70.280, $p < 0.001$; $R^2 = 0.171$) revealed a distinct pattern. Household (C3) and Omnivore (C5) shoppers purchased the most items per transaction ($M = 2.649$ and 2.667, respectively), reflecting the bulk purchase of lower-priced household goods in C3 and comprehensive basket-filling in C5. Minimal shoppers (C1) purchased the fewest items ($M = 1.657$), consistent with their narrower category breadth observed in the class profile.

Overall satisfaction also differed across classes (*Wald* = 21.759, $p < 0.001$; $R^2 = 0.038$). Omnivore (C5; $M = 3.865$) and Auto & Home (C4; $M = 3.738$) shoppers reported the highest satisfaction, while Fashion shoppers (C2; $M = 3.279$) were the least satisfied. A robustness check treating all three shopping behavior variables as ordinal rather than continuous produced substantively identical results (Appendix D).

4.4.2. Platform Awareness Channels

The six platform awareness channels capture how respondents first became aware of the platform and were grouped into three categories according to the initiating agent: active search (consumer-initiated), advertising exposure (firm-initiated), and social information (socially mediated). Because the two consumer-initiated items comprising active search reflect the channel of initial awareness rather than the intensity or extent of subsequent search activity, they are more precisely characterized as active-search-related awareness channels, with active search retained as shorthand throughout. Five of the six channels showed significant class differences.

Both active search channels strongly differentiated the classes. Product searches on the app/web (*Wald* = 23.549, $p < 0.001$; $R^2 = 0.055$) and searching for deals/benefits (*Wald* = 34.753, $p < 0.001$; $R^2 = 0.084$) reached the highest proportions among Omnivore shoppers (C5; 0.728 and 0.638) and the lowest among Fashion (C2; 0.226 and 0.090) and Minimal (C1; 0.277 and 0.136) shoppers. Auto & Home (C4; 0.422 and 0.269) and Household (C3; 0.410 and 0.159) shoppers fell between these extremes, with their proportions in product search approaching twice those of C1 and C2.

Within advertising exposure, the two channels diverged in their ability to differentiate. TV/radio/offline ads showed the second-strongest class differentiation overall (*Wald* = 31.543, $p < 0.001$; $R^2 = 0.080$), with Omnivore shoppers (C5) recording the highest proportion (0.631) and Minimal (C1; 0.107) and Fashion (C2; 0.111) shoppers the lowest. Household shoppers (C3) recorded a notably higher proportion (0.213) than the other three classes (C1: 0.107; C2: 0.111; C4: 0.145). SNS/social media/YouTube ads was the only channel that approached conventional significance, but did not reach it (*Wald* = 9.135, $p = 0.058$; $R^2 = 0.018$). This divergence between the two advertising channels is consistent with the ongoing shift in the advertising landscape from legacy broadcast media toward internet platforms [3]. As audiences migrate away from traditional broadcasting toward internet-based media—a displacement well documented in the Korean market [20]—exposure to traditional channels such as TV, radio, and offline advertising becomes concentrated in an increasingly distinct subset of consumers. This concentration manifests as pronounced class differences, exemplified by the elevated proportion among Household shoppers. Ads on SNS, social media, and YouTube, by contrast, circulate through what has become the mainstream advertising medium, reaching a broader and demographically less differentiated base of users. First awareness through this channel is accordingly spread relatively evenly across the purchase-based segments, compressing class differences below conventional significance.

Both social information channels showed significant yet modest class differentiation. SNS/social media/YouTube reviews (*Wald* = 12.623, $p = 0.013$; $R^2 = 0.027$) and acquaintance recommendations (*Wald* = 11.563, $p = 0.021$; $R^2 = 0.023$) again peaked in Omnivore shoppers (C5; 0.593 and 0.586), with the remaining classes ranging between 0.201 and 0.343.

Collectively, class-level awareness profiles followed a consistent order across the three categories. Omnivore shoppers (C5) scored the highest for every channel, while Minimal (C1) and Fashion (C2) shoppers consistently recorded the lowest proportions. Auto & Home (C4) and Household (C3) shoppers fell between these extremes but stood out specifically for product searches on the app/web, where their proportions (0.422 and 0.410) approached twice those of C1 (0.277) and C2 (0.226). This pattern of stronger active-search behavior distinguishes C3 and C4 from the lower-intensity segments and is the dimension carried forward into the positioning analysis in Section 4.5.

To examine whether awareness channels were directly associated with shopping behavior beyond their shared dependence on class membership, supplementary cross-tabulation analyses were conducted (Appendix E). At the full-sample level, several channels showed

significant but weak associations with shopping behavior (all Cramér's $V \leq 0.162$): searching for deals/benefits, SNS/social media/YouTube reviews, and TV/radio/offline ads were associated with higher per-transaction expenditure, while acquaintance recommendation and TV/radio/offline ads were associated with higher overall satisfaction. Items per transaction showed no significant chi-square association with any channel. Within classes, however, these associations largely dissipated: of the 72 within-class coefficients, only six reached significance (8.3%), close to the rate expected by chance, and they formed no systematic pattern across classes or channels. This pattern indicates that the modest channel–behavior associations observed at the aggregate level operate primarily through class membership, reinforcing the conclusion that segment structure, rather than awareness channel per se, organizes shopping behavior on these platforms.

4.5. Integrated Class Profiles

Combining the dimensions examined above—purchase profiles, demographic covariates, shopping behavior, and platform awareness channels—yields five distinct Korean consumer portraits on Chinese CBEC platforms. Figure 3 presents the integrated profile across all dimensions.

Minimal shoppers (C1; 36.9%), the largest segment, showed a utilitarian purchase profile and a sharply skewed demographic composition: overwhelmingly male (92.3%), moderate-income, and leaning toward AliExpress (63.7%). C1 recorded the lowest per-transaction expenditure and the fewest items per transaction of any class and remained at low-to-moderate levels across all six awareness channels, with no single channel reaching 0.400 (Table 5).

Fashion shoppers (C2; 26.0%), the second-largest segment, concentrated almost exclusively in fashion categories and was female-skewed (62.5%), with the highest share of teenagers and of unmarried respondents of any class, and Temu-leaning in affiliation (63.2%). Its most distinctive behavioral signature was the combination of the lowest active-search proportions and the lowest overall satisfaction of any class (Table 5).

Household shoppers (C3; 17.1%) showed a home-oriented purchase profile and was the most female-skewed (76.1%) as well as the lowest-income class, with a Temu-leaning affiliation (67.4%). Its shopping behavior combined a high number of items per transaction with low per-transaction expenditure—a many-items, modest-spending pattern (Table 5). For platform awareness, C3 showed a distinctive elevation in product searches on app/web (0.410) and in TV/radio/offline ads (0.213); the latter value nearly doubled those recorded by C1 and C2.

Auto & Home shoppers (C4; 16.3%) showed the broadest purchase portfolio after C5 and had a demographic profile that was strongly male (82.5%), predominantly married, and high-income, with the strongest AliExpress affiliation of any class (68.9%). Per-transaction expenditure, items per transaction, and overall satisfaction were all high (Table 5). For platform awareness, C4 was elevated on both active-search channels (0.422 and 0.269), forming, with C3, the second active-search tier beneath C5.

Omnivore shoppers (C5; 3.7%), the smallest segment, purchased across virtually all 16 product categories. This group was gender-balanced (54.2% female), belonged to the highest-income class, and had the strongest Temu affiliation of any segment (86.8%). Per-transaction expenditure, items per transaction, and overall satisfaction all reached their highest values in C5, and its awareness channel proportions were uniformly saturated, at or above 0.586 for every channel (Table 5). Even categories with low base rates in the overall sample recorded high probabilities in this class (Table 3), underscoring a small group of intensive users purchasing across the platform's full assortment.

	C1 Minimal (36.9%)	C2 Fashion (26.0%)	C3 Household (17.1%)	C4 Auto & Home (16.3%)	C5 Omnivore (3.7%)
Purchase profile					
Dominant categories	Tools, electronics, auto	Fashion (clothing + accessories)	Kitchen, household, fashion, stationery	Auto, tools, kitchen, sports, fashion	Virtually all categories
Demographics					
Gender	Strongly male (92.3%)	Female-leaning (62.5%)	Strongly female (76.1%)	Strongly male (82.5%)	Balanced (54.2% female)
Age	Even, 20s–60s (mild 40s)	Youngest (10s: 10.2%)	Even, 30s–60s	40s peak (37.6%)	30s peak (31.7%)
Marital status	Majority married (60.4%)	Most unmarried (43.8%)	Majority married (57.0%)	Most married (76.4%)	Most married, no divorced (67.5%)
Income	Moderate (5.48)	Near moderate (5.23)	Lowest (4.68)	High (6.27)	Highest (6.70)
Platform					
Platform affiliation	AliExpress-leaning (63.7%)	Temu-leaning (63.2%)	Temu-leaning (67.4%)	AliExpress-leaning (68.9%)	Temu-dominant (86.8%)
Shopping behavior					
Per-transaction expenditure	Lowest	Low	Low	High	Highest
Items per transaction	Lowest	Low	High	High	Highest
Overall satisfaction	Moderate	Lowest	Moderate	High	Highest
Active search					
Product search on app/web	Low	Lowest	High	High	Highest
Searching for deals/benefits	Low	Lowest	Low	High	Highest
Advertising exposure					
TV/radio/offline ads	Low	Low	High	Low	Highest
SNS/social media/YouTube ads	Moderate	Moderate	Moderate	Moderate	Highest
Social information					
SNS/social media/YouTube reviews	Low	Low	Lowest	Moderate	Highest
Acquaintance recommendation	Low	Low	Low	Moderate	Highest

Figure 3. Integrated class profiles across the five-class solution. (Cells are shaded by relative standing within each row, from lightest (lowest) to darkest (highest)).

Figure 4 distills these portraits into a positioning map, locating the five classes along the two dimensions that provide the sharpest between-class separation. On the horizontal axis, active-search intensity clearly separates the classes: C2 and C1 at the low end, C3 in the mid-range, C4 above them, and C5 alone in the high-active region. On the vertical axis, platform affiliation distributes the classes along the AliExpress–Temu continuum, with C1 and C4 on the AliExpress-leaning side, and C2, C3, and C5 on the Temu-leaning side. This combination produces a structure in which the two active-search classes (C3 and C4) occupy opposite platform quadrants despite their shared awareness pathway, while C5 occupies the extreme corner on both dimensions.

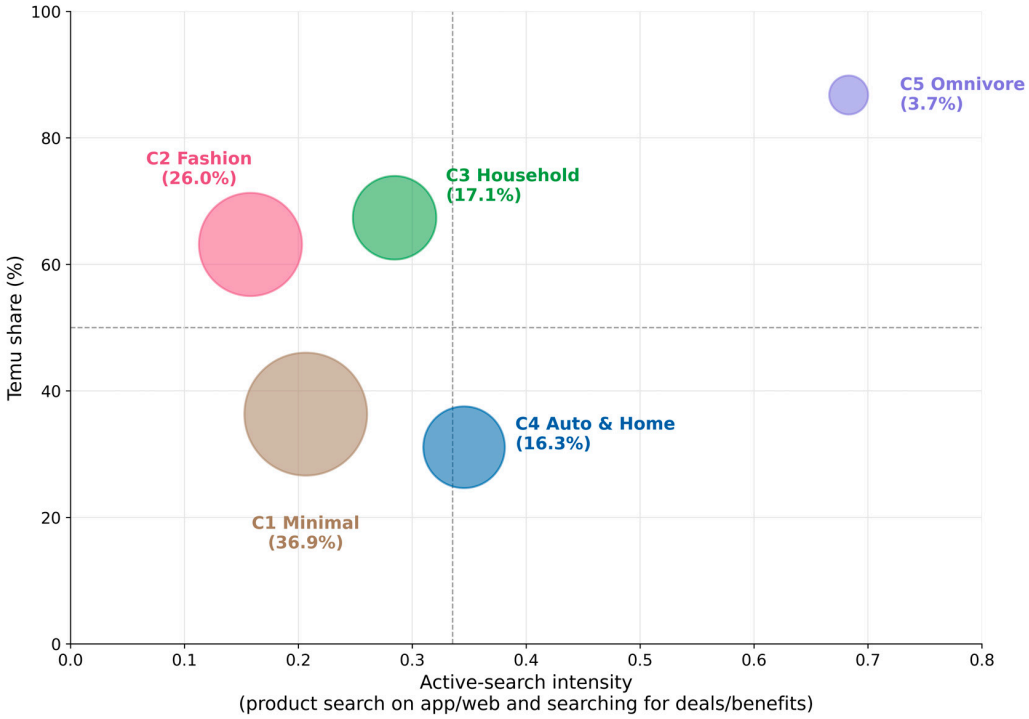


Figure 4. Positioning of the five latent classes by active-search intensity and Temu share. The horizontal axis is active-search intensity (mean of product search on app/web and searching for deals/benefits); the vertical axis is Temu share (%). The bubble areas are proportional to the class sizes. The horizontal reference line marks a 50% platform split, and the vertical reference line marks the across-class mean of active-search intensity (0.336).

5. Discussion

5.1. Summary of Findings

This three-stage analytical framework yielded a systematic picture of consumer heterogeneity on CBEC platforms. LCA of 16 product purchase indicators identified five qualitatively different consumer segments: Minimal shoppers, whose narrow portfolio centered on tools, electronics, and automotive goods; Fashion shoppers, whose purchase pattern concentrates almost exclusively in fashion clothing and accessories; Household shoppers, whose portfolio centered on kitchen and household categories; Auto & Home shoppers, who combined automotive accessories, tools and hardware, and home-related categories with selective fashion purchases; and Omnivore shoppers, who exhibited elevated purchase probabilities across nearly all product categories.

Bias-adjusted three-step analysis (R3STEP) of the demographic and platform covariates revealed that all five examined covariates predicted segment membership. Demographic profiles sharply differed across segments: Minimal and Auto & Home shoppers were predominantly male, Household and Fashion shoppers leaned toward female, and Omnivore shoppers fell between. Age distributions also varied: Auto & Home shoppers were concentrated in their 40s, Omnivore shoppers were concentrated in their 30s, and Fashion shoppers showed the youngest age profile. Marital status patterns paralleled these age distributions, with Auto & Home shoppers showing the highest married share and Fashion shoppers showing the highest unmarried share. Income differentiated segments more modestly: Omnivore and Auto & Home shoppers reported the highest incomes, Household shoppers the lowest, and Minimal and Fashion shoppers fell near the sample mean. Platform affiliation revealed substantively different consumer compositions on the two cross-border platforms: AliExpress users concentrated in the Minimal and Auto & Home

segments, while Temu users concentrated in Fashion, Household, and most distinctively, Omnivore segments, where 86 percent of members used Temu.

The BCH distal outcome analysis of shopping behavior showed that the segments systematically differed in per-transaction expenditure, items per transaction, and overall satisfaction. Omnivore shoppers spent the most per transaction, followed by Auto & Home shoppers, while Minimal and Fashion shoppers spent the least. Items per transaction reflected a distinct pattern: Household and Omnivore shoppers purchased the most items per transaction, while Minimal shoppers purchased the least. Overall satisfaction differed more modestly, with Omnivore and Auto & Home shoppers reporting the highest satisfaction and Fashion shoppers the lowest.

A parallel BCH analysis of platform awareness channels further revealed a structured pattern of consumer-side variation across channel types. Omnivore shoppers exhibited elevated reception across all awareness channels. Auto & Home shoppers occupied the second-highest position on most channels and distinguished themselves in active searches. Household shoppers showed a selective elevation in product searches on app/web and on TV/radio/offline ads. Fashion and Minimal shoppers relied on substantially narrower channel sets, with Fashion shoppers showing the lowest overall use of active search channels.

5.2. Theoretical Implications

The findings yield three theoretical implications for interactive marketing research on CBEC platforms. First, consumer populations on CBEC platforms exhibit qualitatively distinct subpopulation structures that aggregate frameworks overlook. Five qualitatively different consumer segments emerge from the present analysis—differing not in degree but in the cross-category purchase portfolios they bring to the platform—directly demonstrating that the homogeneity assumption embedded in prior CBEC consumer research is empirically inadequate. The methodological foundation for this contribution rests on the established premise that variable-centered approaches, by treating the population as a single underlying group, cannot reveal subpopulations whose underlying parameters differ from one another [17]. This contribution builds on earlier work in interactive marketing research, where the latent class segmentation of a major retailer's digital music customer base revealed segment-level patterns that aggregate-level analyses could not systematically detect [13]. The same analytical orientation continues to uncover a meaningful subpopulation structure in more recent work, with applications including the segmentation of individual adoption of new consumer packaged goods [14], confirming that the heterogeneity-segmentation orientation remains an active research stream rather than a historical curiosity. The present study extends this orientation to CBEC, a setting in which platform-orchestrated value creation has been conceptually established, but consumer-side variation has remained empirically under-examined. The implication is that interactive marketing research on platform contexts benefits from person-centered methods that render visible subpopulation structures beyond the reach of aggregate frameworks.

Second, cross-category purchase behavior, properly operationalized for contemporary CBEC settings, is a substantively informative segmentation basis. It captures portfolio-level consumer heterogeneity that attitudinal and perceptual segmentation bases cannot reveal directly. The five segments identified through the present analysis qualitatively reflect the different cross-category purchase patterns reported across the 16 product categories. The original validity of cross-category purchase behavior as a segmentation basis was established in early-web-era online retailing, where seven product categories—dominated by computer hardware, software, and home electronics, with these three split along a price dimension—yielded segments distinguishable by benefits sought [8]. Subsequent latent class applications in e-commerce have largely shifted away from cross-category

purchase indicators, instead segmenting consumers on consumer attitudes, perceptions, or psychographic dispositions such as channel attitudes [7], perceived channel value differences [19], psychographic factors [6], and constructs from the unified theory of acceptance and use of technology (UTAUT2) augmented with trust [21]. The present analysis revisits cross-category purchase behavior as the segmentation basis but with a sixteen product-category-level indicator set that reflects the broader consumer goods spectrum supplied by contemporary CBEC platforms. The five-segment structure that emerges could not have been detected through attitudinal or perceptual bases: portfolio-level heterogeneity—who buys fashion only, who buys household goods primarily, who buys across nearly all categories—is a structural feature of consumer purchase behavior that attitudinal and perceptual constructs do not capture directly. As e-commerce platforms have evolved from technology-dominated retailing toward a broader consumer-goods supply, segmentation bases must evolve correspondingly, and behavioral indicators offer the empirical leverage that contemporary platform contexts require.

Third, active search behavior distinguishes consumer segments on CBEC platforms as a dimension in which qualitatively distinct segments are systematically positioned. Active search is not merely a continuous variable that varies across an undifferentiated consumer population; the present analysis demonstrates that segments occupy distinct positions on this dimension, with Omnivore shoppers exhibiting uniformly elevated active search across the relevant channels, Auto & Home shoppers occupying the second-highest position with notable elevation in product search and search for deals/benefits, Household shoppers showing selective elevation in product search on app/web, and Fashion and Minimal shoppers relying on substantially narrower search effort. The theoretical foundation for treating active search as a coherent segmentation-relevant dimension relies on the established characterization of consumer search as a behavior whose magnitude systematically varies across consumers as a function of individual-level factors [23]. A subsequent integrative framework organized the determinants of external information search into four mediating constructs: ability, motivation, costs, and benefits [9]. The substantive premise is that consumer-side variation in active search reflects underlying individual-level differences and that segmentation analysis should reveal these differences as identifiable patterns across consumer types. Recent platform-level evidence confirms that consumer search heterogeneity is substantial and consequential on Chinese e-commerce platforms, where naïve and sophisticated consumers differ in search breadth and depth in ways that shape platform-side search design [10]. The present study advances the interactive marketing literature by demonstrating that consumer search heterogeneity on CBEC platforms organizes into a richer segment-level structure than the naïve-sophisticated dichotomy alone reveals: five qualitatively distinct consumer types differ in active search not only in degree but also in the broader purchase portfolio and platform awareness channel patterns they bring to the platform. The theoretical implication is that active search must be treated as a segment-level dimension in interactive marketing research on CBEC platforms, with strategic and analytical consequences that aggregate-level treatments cannot deliver.

5.3. Practical Implications

This study's practical implications are threefold. The first concerns Household shoppers (C3, 17.1%), a predominantly female and low-income segment whose purchases concentrate in the household and kitchen categories, characterized by high-volume, low-unit-price buying. Two characteristics of this segment depart from the conventions of CBEC marketing. Reception of digital social channels is the lowest in the sample for this segment, while reception of traditional broadcast advertising is second only to that of Omnivore shoppers, suggesting that offline media should not be overlooked in the channel mix for

this segment. Their purchasing pattern—many items with modest total spending—also calls for promotional structures that reward purchasing more items per transaction, such as buy-more-save-more discounts on household-category bundles, rather than premium positioning or single-item discounting. Thus, these purchasing and reception patterns can be construed as support for bundle-oriented promotions in their core household and kitchen categories, combined with a continued role for offline media in the channel mix.

A second implication concerns AliExpress's predominantly male user base, which is not a single cohort but comprises two distinct segments: Minimal shoppers (C1, 36.9%) and Auto & Home shoppers (C4, 16.3%). These segments differ qualitatively in portfolio, expenditure, and active search, and therefore call for differentiated promotions and recommendation-system tuning rather than uniform male-targeted communications. Minimal shoppers exhibit a narrow utility-oriented portfolio centered on tools, electronics, and automotive accessories, characterized by low expenditure and minimal active search. For them, recommendations and promotions could focus on single products with accessible price points within these utility categories.

Conversely, Auto & Home shoppers exhibit a broader portfolio spanning automotive accessories, tools, kitchen, sports, and furniture, characterized by high expenditure, active search behavior, and a notable increase in deal-and-benefit searches. For them, recommendations and promotions could involve portfolio bundles and cross-category suggestions that leverage their willingness to search across the platform's full inventory. Thus, these segment differences suggest moving beyond uniform male-targeted communications toward promotions tailored separately to the utility-narrow Minimal type and the deal-active, broad-portfolio Auto & Home type.

The third implication concerns Fashion shoppers (C2, 26.0%), a predominantly female segment that is the youngest and most single-leaning, with a portfolio concentrated almost exclusively in fashion clothing and accessories on Temu. The most notable pattern is that this segment concentrates its purchasing most heavily in a single category while reporting the lowest overall satisfaction in the sample. This coexistence of heavy purchasing and low satisfaction suggests a gap between what draws these consumers to the fashion offering and what the shopping experience ultimately provides: the category's value proposition attracts purchases from its core segment more effectively than it delivers on its promise. Thus, for fashion merchants, the practical task this pattern identifies is to close that gap by directing improvement efforts toward the fashion-category shopping experience itself.

5.4. Limitations and Future Research

This study draws on a single-country sample of Korean CBEC users, which limits the generalizability of the five-segment structure beyond the Korean market. CBEC penetration and category-portfolio preferences may substantially differ across national markets, with the Korean context potentially reflecting features specific to its well-developed domestic e-commerce ecosystem. Korean consumers with extensive e-commerce experience may approach cross-border platforms differently from novice users encountering online retail for the first time. They may exhibit more deliberate, category-specialized purchase patterns than the exploratory, cross-platform usage that less experienced consumers might display. Future research applying the present analytical framework to other national contexts, particularly those at different stages of CBEC maturity, would distinguish generalizable features of the segment structure from context-specific ones.

The second limitation concerns the cross-sectional design. The five identified segments are estimated from a single point in time, leaving open the question of whether the segment composition remains stable or changes over time. As individual consumers' cross-border purchasing experience deepens, they may transition between segments, with the resulting

effects manifesting as shifts in segment proportions and changes in segment boundaries. For instance, the persistently dissatisfied Fashion shoppers identified here may either migrate toward the broader-portfolio Omnivore pattern as they gain cross-border platform experience or leave the cross-border ecosystem entirely as their dissatisfaction accumulates—two trajectories that carry sharply different implications for platform retention strategies. Future research using longitudinal or repeated cross-sectional designs could examine such temporal stability of the segment structure and trajectories of consumer migration between segments, potentially through latent growth curve analysis [14].

The third limitation concerns the underlying mechanisms producing the observed segment heterogeneity. This analysis is grounded in cross-category purchase behavior and characterizes how the five segments differ in observable behavioral and platform awareness outcomes. However, it does not examine the psychological mechanism that plausibly underlies these differences. Without measures of dispositional and motivational factors, the study cannot distinguish among competing accounts of why behavioral patterns sharply differ across segments—whether such differences originate from distinct value perceptions, motivational orientations, or trait-level individual differences in shopping behavior. Future research integrating behavioral indicators with psychological measures would permit a direct test of whether segmentations based on behavioral and psychological constructs yield the same consumer typology or reveal distinct subpopulation structures.

6. Conclusions

This study identifies five qualitatively distinct consumer segments on Korean CBEC platforms— Minimal shoppers, Fashion shoppers, Household shoppers, Auto & Home shoppers, and Omnivore shoppers—through bias-adjusted LCA of cross-category purchase patterns. The integrated analytical procedure—LCA of 16 product purchase indicators, three-step covariate analysis (R3STEP) of demographic and platform-affiliation predictors, and bias-adjusted distal-outcome analysis (BCH) of shopping behavior and platform awareness channels—yields a unified profile of each segment. These findings have a broader implication: consumer populations on CBEC platforms exhibit subpopulation structures that aggregate frameworks overlook. Therefore, interactive marketing research on cross-border platforms requires analytical commitments capable of revealing the structural variations that aggregate methods cannot detect.

Funding: This work was supported by a Research Grant of Pukyong National University (2025).

Institutional Review Board Statement: Ethical review and approval were waived for this study by the Institutional Review Board of Pukyong National University (PKNU IRB; File No. PKNU IRB 2026-03-013), confirmed on 9 April 2026, on the grounds that the research involved the secondary use of existing, de-identified data and documents pertaining to research participants, with no direct contact between the investigator and human subjects.

Informed Consent Statement: Not applicable. The study used de-identified secondary data with no direct contact between the investigator and research subjects.

Data Availability Statement: Restrictions apply to the availability of these data. Data were obtained from Opensurvey (https://www.opensurvey.co.kr/, accessed on 3 October 2024) and are available from Opensurvey with their permission, subject to licensing arrangements.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A. Model Comparison for Latent Class Enumeration (1- to 10-Class Models Without Direct Effects)

Classes	LL	BIC	AIC	AIC3	Entropy R^2	Class. Error	Max BVR	Bootstrap p	−2LL Diff	BLRT p	Smallest (%)
1	−4870.698	9843.747	9773.396	9789.396	1.000	0.000	188.306	—	—	—	100.000
2	−4448.578	9108.254	8963.156	8996.156	0.895	0.013	49.675	<0.001	844.240	<0.001	9.510
3	−4329.327	8978.501	8758.654	8808.654	0.733	0.072	40.712	0.052	238.502	<0.001	3.720
4	**−4256.938**	**8942.471**	**8647.877**	**8714.877**	**0.679**	**0.155**	**11.438**	**0.108**	**144.778**	**<0.001**	**3.690**
5	−4219.251	8975.844	8606.502	8690.502	0.693	0.165	5.410	0.180	75.374	<0.001	3.680
6	−4195.052	9036.194	8592.104	8693.104	0.686	0.191	4.752	0.114	48.398	0.002	3.660
7	−4172.029	9098.896	8580.059	8698.059	0.694	0.191	3.445	0.102	46.046	0.006	2.160
8	−4151.640	9166.865	8573.279	8708.279	0.714	0.181	3.829	0.072	40.779	0.014	1.150
9	−4132.810	9237.954	8569.620	8721.620	0.720	0.193	2.750	0.078	37.659	0.046	2.860
10	−4117.965	9317.011	8573.930	8742.930	0.734	0.192	3.190	0.060	29.691	0.190	2.220

Note. n = 600. No direct effects. Bold font indicates the model selected for direct effects specification. Class. error = Classification error. BLRT = Bootstrap likelihood-ratio test (500 replications).

Appendix B. Direct Effects Freed in the Five-Class Model

Pair	Item 1	Item 2	BVR
1	Fashion clothing	Fashion accessories	5.007
2	Cosmetics/perfume	Pet supplies	4.684
3	Furniture/interior	Stationery/hobby goods	5.410
4	Tools/hardware	Books/media	5.082
5	Pet supplies	Books/media	4.131

Note. BVR = bivariate residual from the five-class model without direct effects. Pairs with BVR > 3.840 were freed as direct effects.

Appendix C. Bivariate Residual Diagnostics for the Final Five-Class Solution (With Five Direct Effects Freed)

Indicator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Fashion clothing	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2. Fashion accessories	0.000	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
3. Cosmetics/perfume	1.302	0.814	—	—	—	—	—	—	—	—	—	—	—	—	—	—
4. Sports/leisure	0.657	0.002	0.150	—	—	—	—	—	—	—	—	—	—	—	—	—
5. Household essentials	0.484	3.271	0.466	0.109	—	—	—	—	—	—	—	—	—	—	—	—
6. Kitchen supplies	0.334	0.045	0.382	0.017	0.162	—	—	—	—	—	—	—	—	—	—	—
7. Maternity/childcare	1.131	0.095	0.239	1.301	0.331	0.576	—	—	—	—	—	—	—	—	—	—
8. Furniture/interior	0.023	0.614	0.594	0.689	0.158	1.921	2.647	—	—	—	—	—	—	—	—	—
9. Electronics/digital	0.236	1.739	0.036	0.002	0.014	0.335	0.002	0.376	—	—	—	—	—	—	—	—
10. Food/groceries	0.170	1.716	1.101	0.057	0.021	0.218	0.146	1.703	2.870	—	—	—	—	—	—	—
11. Automotive accessories	0.001	1.236	0.241	0.007	0.014	0.020	0.008	0.206	0.001	0.221	—	—	—	—	—	—
12. Tools/hardware	0.006	0.042	0.626	0.308	2.884	0.430	1.021	1.315	0.291	0.323	1.535	—	—	—	—	—
13. Pet supplies	0.046	0.324	0.000	1.095	1.086	0.168	0.683	0.542	0.188	2.275	0.029	0.369	—	—	—	—
14. Health/medical	0.072	0.805	0.670	0.176	0.104	0.073	0.398	0.278	0.354	0.121	0.058	0.450	0.513	—	—	—
15. Stationery/hobby	0.068	0.667	3.585	0.100	0.438	0.034	0.009	0.000	0.009	0.323	0.089	0.249	0.003	0.027	—	—
16. Books/media	0.498	0.153	0.669	0.380	0.233	0.001	2.632	0.404	0.676	0.124	0.046	0.000	0.000	0.981	1.211	—

Note. Column numbers correspond to the indicator numbers in the first column. Maximum BVR = 3.585 (<3.840 threshold). Direct effects freed: Fashion clothing × Fashion accessories, Cosmetics/perfume × Pet supplies, Furniture/interior × Stationery/hobby goods, Tools/hardware × Books/media, Pet supplies × Books/media.

Appendix D. BCH Distal Outcome Analysis: Shopping Behavior with Ordinal Treatment

Outcome	C1	C2	C3	C4	C5	Wald	p	R^2
	Minimal Shoppers	**Fashion Shoppers**	**Household Shoppers**	**Auto & Home Shoppers**	**Omnivore Shoppers**			
Shopping behavior (BCH-adjusted means, ordinal treatment)								
Per-transaction expenditure	1.571	1.794	1.808	2.130	3.404	47.465	<0.001	0.130
Items per transaction	1.657	1.944	2.648	2.345	2.661	35.633	<0.001	0.171
Overall satisfaction	3.489	3.279	3.458	3.738	3.862	18.307	0.001	0.038

Note. Robustness check with ordinal treatment. See Table 5 for the primary analysis with continuous treatment. Bold font indicates the outcome domain.

Appendix E. Associations Between Platform Awareness Channels and Shopping Behavior

Table A1. Full-sample associations (n = 600).

Awareness Channel	Per-Transaction Expenditure			Items per Transaction			Overall Satisfaction		
	χ^2	V	τ	χ^2	V	τ	χ^2	V	τ
Active search									
Product search on app/web	4.354	0.085	0.078 *	3.790	0.079	0.069	0.308	0.023	−0.004
Searching for deals/benefits	15.707 **	0.162	0.127 **	1.897	0.056	0.041	1.982	0.057	0.039
Social information									
Acquaintance recommendation	6.044	0.100	0.089 *	3.739	0.079	−0.043	14.122 **	0.153	0.129 **
SNS/social media/YouTube reviews	13.705 **	0.151	0.118 **	0.226	0.019	−0.009	7.608	0.113	0.058
Advertising exposure									
SNS/social media/YouTube ads	3.924	0.081	0.060	7.236	0.110	0.083 *	4.101	0.083	0.031
TV/radio/offline ads	14.899 **	0.158	0.128 **	4.468	0.086	0.056	9.518 *	0.126	0.115 **

Note. χ^2 = Pearson chi-square; V = Cramér's V; τ = Kendall's tau-b. For chi-square tests, per-transaction expenditure and items per transaction were collapsed at the top (four or more categories merged) to satisfy expected cell count requirements (df = 3); overall satisfaction retained its original five categories (df = 4). Kendall's tau-b was computed on the original uncollapsed scales. * $p < 0.05$, ** $p < 0.01$.

Table A2. Within-class associations, Kendall's tau-b (modal class assignment).

Outcome	Awareness Channel	C1	C2	C3	C4
		Minimal Shoppers (n = 228)	**Fashion Shoppers (n = 162)**	**Household Shoppers (n = 96)**	**Auto & Home Shoppers (n = 92)**
Per-transaction expenditure	Product search on app/web	0.079	−0.060	0.141	0.049
	Searching for deals/benefits	0.005	0.124	0.068	0.060
	Acquaintance recommendation	0.008	0.014	0.125	0.113
	SNS/social media/YouTube reviews	0.048	0.074	0.126	0.216 *
	SNS/social media/YouTube ads	−0.043	0.128	0.005	−0.070
	TV/radio/offline ads	0.056	−0.007	0.022	0.149
Items per transaction	Product search on app/web	0.105	−0.011	0.135	0.123
	Searching for deals/benefits	0.027	−0.035	−0.021	0.083
	Acquaintance recommendation	−0.131 *	−0.088	0.155	−0.051
	SNS/social media/YouTube reviews	−0.043	0.051	−0.107	−0.007
	SNS/social media/YouTube ads	0.002	0.193 **	0.029	−0.089
	TV/radio/offline ads	0.013	0.031	−.138	0.073
Overall satisfaction	Product search on app/web	−0.082	−0.007	−0.045	0.167
	Searching for deals/benefits	−0.003	−0.009	0.051	0.004
	Acquaintance recommendation	0.086	0.201 **	0.054	0.148
	SNS/social media/YouTube reviews	0.147 *	−0.023	0.022	−0.016
	SNS/social media/YouTube ads	−0.065	0.118	0.009	−0.079
	TV/radio/offline ads	0.127 *	0.001	−0.029	0.161

Note. Class sizes reflect modal assignment; model-estimated class proportions appear in Table 2. Class 5 (Omnivore shoppers, n = 22) was excluded from segment-specific testing owing to insufficient cell sizes for valid inference. Within-class chi-square tests are not reported because expected cell counts fell below conventional thresholds in most tables; Kendall's tau-b does not depend on expected cell counts. Of the 72 within-class coefficients, six reached significance (8.3%), close to the rate expected by chance. * $p < 0.05$, ** $p < 0.01$.

References

1. Min, Y. AliExpress and Temu Earned over 4 Trillion Won in South Korea Last Year, with Average Payment of 80,000 Won per Person. *Chosun Biz*, 11 February 2025. Available online: https://biz.chosun.com/distribution/channel/2025/02/11/RJLBNEBKAZEMPC6FCJBZ323XDI/(accessed on 10 April 2026).
2. Liu, C.; Chen, T.; Pu, Q.; Jin, Y. Text Mining for Consumers' Sentiment Tendency and Strategies for Promoting Cross-Border E-Commerce Marketing Using Consumers' Online Review Data. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 125. [CrossRef]
3. Wang, C.L. New frontiers and future directions in interactive marketing: Inaugural editorial. *J. Res. Interact. Mark.* **2021**, *15*, 1–9. [CrossRef]
4. Wang, C.; Liu, T.; Zhu, Y.; Wang, H.; Wang, X.; Zhao, S. The influence of consumer perception on purchase intention: Evidence from cross-border e-commerce platforms. *Heliyon* **2023**, *9*, e21617. [CrossRef] [PubMed]
5. Xu, Y.; He, D.; Fan, M. Antecedent research on cross-border e-commerce consumer purchase decision-making: The moderating role of platform-recommended advertisement characteristics. *Heliyon* **2024**, *10*, e37627. [CrossRef] [PubMed]
6. Huseynov, F.; Özkan Yıldırım, S. Online consumer typologies and their shopping behaviors in B2C e-commerce platforms. *SAGE Open* **2019**, *9*, 2158244019854639. [CrossRef]
7. Konuş, U.; Verhoef, P.C.; Neslin, S.A. Multichannel shopper segments and their covariates. *J. Retail.* **2008**, *84*, 398–413. [CrossRef]
8. Bhatnagar, A.; Ghose, S. A latent class segmentation analysis of e-shoppers. *J. Bus. Res.* **2004**, *57*, 758–767. [CrossRef]
9. Schmidt, J.B.; Spreng, R.A. A proposed model of external consumer information search. *J. Acad. Mark. Sci.* **1996**, *24*, 246–256. [CrossRef]
10. Li, J.; He, C.; Shi, Z. Platform search design with heterogeneous consumers. *Humanit. Soc. Sci. Commun.* **2025**, *12*, 580. [CrossRef]
11. Deighton, J.; Sorrell, M. The future of interactive marketing. *Harv. Bus. Rev.* **1996**, *74*, 151–160.
12. Ayeni, A.A.; Bisschoff, C.A.; Sekhampu, J.; Lekunze, J.; Ayeni, E. Interactive marketing: A systematic review and research agenda. *F1000Research* **2024**, *13*, 1383. [CrossRef]
13. Reimer, K.; Rutz, O.J.; Pauwels, K. How online consumer segments differ in long-term marketing effectiveness. *J. Interact. Mark.* **2014**, *28*, 271–284. [CrossRef]
14. Allaway, A.W.; McCullough, H.; Kim, K.; McAmis, G. Segmenting customer adoption of new consumer packaged goods using latent growth curve modeling. *J. Mark. Theory Pract.* **2024**, *32*, 25–42. [CrossRef]
15. Block, J. *Lives Through Time*; Bancroft Books: Berkeley, CA, USA, 1971.
16. Laursen, B.; Hoff, E. Person-centered and variable-centered approaches to longitudinal data. *Merrill-Palmer Q.* **2006**, *52*, 377–389. [CrossRef]
17. Howard, M.C.; Hoffman, M.E. Variable-centered, person-centered, and person-specific approaches: Where theory meets the method. *Organ. Res. Methods* **2018**, *21*, 846–876. [CrossRef]
18. Magidson, J.; Vermunt, J.K. Latent class models for clustering: A comparison with K-means. *Can. J. Mark. Res.* **2002**, *20*, 36–43.
19. Wang, Q.; Yang, X.; Song, P.; Sia, C.L. Consumer segmentation analysis of multichannel and multistage consumption: A latent class MNL approach. *J. Electron. Commer. Res.* **2014**, *15*, 339–352. Available online: http://www.jecr.org/node/453 (accessed on 10 April 2026).
20. Shim, D.; Lee, C.; Oh, I. Analysis of OTT Users' Watching Behavior for Identifying a Profitable Niche: Latent Class Regression Approach. *J. Theor. Appl. Electron. Commer. Res.* **2022**, *17*, 1564–1580. [CrossRef]
21. Nadeem, W.; Juntunen, M.; Juntunen, J. Consumer segments in social commerce: A latent class approach. *J. Consum. Behav.* **2017**, *16*, 279–292. [CrossRef]
22. Vermunt, J.K. Latent class modeling with covariates: Two improved three-step approaches. *Political Anal.* **2010**, *18*, 450–469. [CrossRef]
23. Punj, G.N.; Staelin, R. A model of consumer information search behavior for new automobiles. *J. Consum. Res.* **1983**, *9*, 366–380. [CrossRef]
24. Bloch, P.H.; Sherrell, D.L.; Ridgway, N.M. Consumer search: An extended framework. *J. Consum. Res.* **1986**, *13*, 119–126. [CrossRef] [PubMed]
25. Honka, E.; Seiler, S.; Ursu, R. Consumer Search: What Can We Learn from Pre-Purchase Data? *J. Retail.* **2024**, *100*, 114–129. [CrossRef]
26. Ursu, R.M.; Zhang, Q.; Honka, E. Search gaps and consumer fatigue. *Mark. Sci.* **2023**, *42*, 110–136. [CrossRef]
27. Amano, T.; Rhodes, A.; Seiler, S. *Flexible Demand Estimation with Search Data*; CEPR Discussion Paper No. 16933; Centre for Economic Policy Research: London, UK, 2022. [CrossRef]
28. Morozov, I.; Seiler, S.; Dong, X.; Hou, L. Estimation of preference heterogeneity in markets with costly search. *Mark. Sci.* **2021**, *40*, 871–899. [CrossRef]
29. Honka, E.; Hortaçsu, A.; Vitorino, M.A. Advertising, consumer awareness, and choice: Evidence from the U.S. banking industry. *RAND J. Econ.* **2017**, *48*, 611–646. [CrossRef]

30. Vermunt, J.K.; Magidson, J. *Upgrade Manual for Latent GOLD Basic, Advanced, Syntax, and Choice Versions 6.0 and 6.1*; Statistical Innovations: Arlington, MA, USA, 2025. Available online: https://www.statisticalinnovations.com/wp-content/uploads/LG6061manual.pdf (accessed on 10 April 2026).
31. Bakk, Z.; Tekle, F.B.; Vermunt, J.K. Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociol. Methodol.* **2013**, *43*, 272–311. [CrossRef]
32. Bolck, A.; Croon, M.; Hagenaars, J. Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Anal.* **2004**, *12*, 3–27. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.