International Journal of Asian Social Science
ISSN(e): 2224-4441
ISSN(p): 2226-5139
DOI: 10.55493/5007.v16i4.5960
Vol. 16, No. 4, 224-239

URL: www.aessweb.com

Accessibility minutes and low-rating risk: Evidence from Malaysian hotel reviews with airport and transit proximity indicators

Zhang Juan[1+]
Choo Wei Chong[2]
Yee Choy Leong[3]
Lin Yihuan[4]

[1]*Institute for Mathematical Research (INSPEM), Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.*
[1]*Email: gs66172@student.upm.edu.my*
[2,3,4]*School of Business and Economics, Universiti Putra Malaysia, 43400 Serdang, Malaysia.*
[2]*Email: wchoo@upm.edu.my*
[3]*Email: fyee@upm.edu.my*
[4]*Email: gs58501@student.upm.edu.my*

(+ Corresponding author)

Article History
Received: 28 January 2026
Revised: 3 March 2026
Accepted: 19 March 2026
Published: 8 April 2026

Keywords
Accessibility measurement
Hotel ratings
Low-rating risk
Mixed-effects logistic regression
Monitoring threshold
Online hotel reviews
Threshold-based governance
Transport nodes
Travel friction
Travel-purpose heterogeneity.

ABSTRACT

Tourism recovery in Malaysia is accelerating ahead of a national tourism campaign, but destination managers and hotels have limited capacity to monitor where arrival frictions translate into reputational risk. This study tests whether objectively measured accessibility is associated with low-rating risk and whether the association differs by travel purpose, with the goal of deriving an operational monitoring threshold. We analyzed 346,988 Ctrip hotel reviews for Malaysia (October 2022–December 2025) linked to 792 hotels. Proximity descriptors were parsed to estimate minutes to the nearest transport node and to classify the nearest node as an airport or transit hub. Low ratings (rating < 3) were modeled using mixed-effects logistic regression with hotel random intercepts, controlling for review length and hotel attributes, complemented by nonlinear minutes specifications and probability-scale predictions. Relative to 0–10 minutes, low-rating odds were lower at 10–20 minutes (OR 0.80, 95% CI 0.71–0.89) and 20–40 minutes (0.73, 0.64–0.82), while >40 minutes was not distinguishable from baseline after adjustment. Business stays showed a higher baseline risk and a widening probability gap as friction increased. Overall, results suggest a practical monitoring trigger around 40 minutes, supporting targeted arrival information, transfer coordination, and reliability buffers where accessibility frictions and business demand coincide.

Contribution/ Originality: This study operationalises platform proximity descriptors as an auditable minutes-to-node accessibility measure and estimates its association with low-rating risk using hierarchical logistic models and probability-scale translations. It advances a threshold-based monitoring logic and documents travel-purpose heterogeneity, strengthening the interpretability and policy relevance of review-derived risk indicators.

1. INTRODUCTION

Malaysia is entering a renewed phase of tourism expansion as the country advances the Visit Malaysia 2026 (VM2026) campaign and its associated calendar of events, which positions tourism as a national growth lever and foregrounds coordination across destinations, operators, and information intermediaries (Ministry of Tourism Arts and Culture, 2026; Tourism Malaysia, 2026). In this policy window, the practical problem is no longer only how to attract visitors, but also how to detect where experience risks are likely to concentrate and how to prioritise mitigation when monitoring capacity is limited. This question is especially salient when destination systems are expected to scale quickly, because quality failures are rarely evenly distributed and therefore require triage rather than uniform intervention.

A persistent yet understudied friction point in destination experiences is arrival and access. Even when the hotel product is stable, travellers still absorb transfer burden, time uncertainty, and last-mile coordination costs, and these frictions are often incorporated into what guests later describe as the “stay” rather than treated as externalities. In platform-mediated markets, such frictions can be amplified because evaluation is filtered through online ratings that directly shape reputation and choice. A robust behavioural mechanism is that negative information tends to be perceived as especially diagnostic, meaning that low-end signals can exert disproportionate influence relative to equally strong positive information (Qahri-Saremi & Montazemi, 2023). Moreover, hotel choice evidence suggests that ratings and review signals can shape consideration, which makes very low ratings a practically consequential risk indicator rather than a marginal outcome (Gavilan, Avello, & Martinez-Navarro, 2018). In hotel settings, online reviews have also been shown to shift consumers’ consideration sets, reinforcing the relevance of low-end rating risk as a governance-relevant signal rather than a purely marginal evaluation (Vermeulen & Seegers, 2009). In other words, the low-rating tail is governance-relevant because it concentrates punitive evaluations that can trigger avoidance, complaint diffusion, and resource-intensive response.

Despite this need, much of the hotel evaluation literature still operationalises “location” using subjective location scores, coarse administrative units, or broad neighbourhood indicators. These measures often conflate structural accessibility with general neighbourhood satisfaction and therefore remain difficult to translate into implementable monitoring rules. This limitation becomes more salient given the broader policy turn toward decision-ready indicators and governance routines rather than diffuse descriptive reporting, a direction emphasised in international tourism governance frameworks and measurement initiatives (OECD, 2024; UN Tourism, 2024). At the same time, platforms increasingly provide semi-structured contextual fields alongside review text. If such fields can be converted into auditable, objective measures, they can complement text mining by producing governance-grade signals that are easier to implement under capacity constraints.

This study addresses that gap by constructing an objective accessibility indicator from platform-provided proximity descriptors and testing how it relates to low-end rating risk in Malaysian hotels. Using Ctrip hotel reviews for Malaysia from October 2022 to December 2025, the paper extracts an estimated travel time in minutes to the nearest transport node and distinguishes whether the nearest node is airport-related or transit-related. The choice of Ctrip is analytically meaningful because it is a major Chinese-language travel platform within Trip.com Group’s ecosystem and therefore provides a window into how Chinese outbound travellers encode and penalise frictions in overseas stays (Ctrip.com, 1999; Trip.com Group Limited, 2024). This positioning matters for governance because cross-border evaluation standards and expectation alignment can shape which frictions are interpreted as “service failures,” and negative signals are known to be especially diagnostic in such judgment environments (Qahri-Saremi & Montazemi, 2023).

The conceptual logic is summarised in Figure 1, which links the VM2026 governance window to an objective minutes-based accessibility signal, then to arrival friction, and finally to a punitive evaluation outcome operationalised as LowRating (rating below 3). The framework also anticipates heterogeneity by travel purpose and motivates probability-scale translation so that model outputs can be expressed as interpretable risk differences rather than only coefficient changes, consistent with contemporary guidance on interpreting nonlinear and moderated models through predicted quantities (Arel-Bundock, Greifer, & Heiss, 2024). The study is designed to advance an implementable risk-screening logic rather than a causal claim about transport infrastructure, and it therefore emphasises monitoring triggers that can be enacted under capacity constraints.

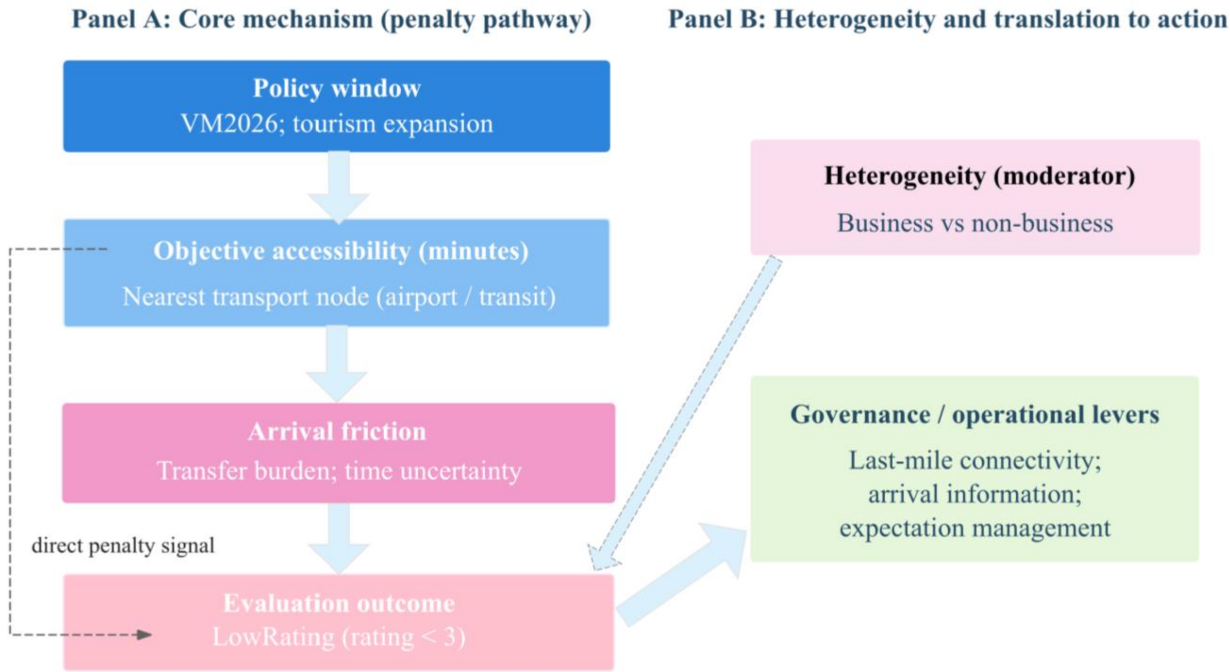

Figure 1. Conceptual framework.

Note: Panel A links the VM2026 policy window to objective accessibility (Minutes to the nearest airport/Transit node), arrival friction, and low-rating risk (Rating < 3). Panel B summarises travel-purpose heterogeneity (Business vs non-business) and the operational levers for mitigating arrival friction.

1.1. Research Gap

Prior work indicates that transport conditions and spatial context can shape accommodation market outcomes, supporting the view that accessibility is a structurally meaningful factor rather than a neutral background attribute (Zheng, Gao, & Fang, 2022). However, empirical evidence remains limited on whether a platform-native, objectively measured accessibility signal can be translated into an operational threshold for low-end rating risk, and whether this translation varies by travel purpose, given documented differences between business and leisure travellers in online hotel evaluations (Kim, Ma, & Park, 2023; Schiessl, Manosso, Alves, & Prado, 2026). The missing link is translational as much as empirical, because the literature has not yet provided a decision-ready accessibility indicator that is auditable, interpretable, and implementable as a monitoring routine under capacity constraints in a destination governance window that increasingly emphasises indicator-based decision support (OECD, 2024; UN Tourism, 2024).

1.2. Research Questions

RQ1 examines whether objective accessibility, measured as minutes to the nearest transport node, is associated with low-rating risk.

RQ2 tests whether travel purpose moderates the accessibility–low-rating association, contrasting business and non-business stays.

RQ3 evaluates whether accessibility can be translated into an operational, probability-scale threshold for monitoring and prioritisation.

1.3. Contributions and Paper Organisation

The study contributes in three aligned ways. First, it offers a measurement contribution by demonstrating how platform proximity descriptors can be converted into an auditable minutes-to-node indicator that is replicable without specialised GIS infrastructure. Second, it provides empirical evidence on how this objective accessibility signal relates to low-rating risk and how the relationship differs by travel purpose. Third, it advances a translational contribution by expressing model outputs in probability space and framing an operational threshold logic that can support capacity-constrained monitoring, consistent with the broader shift toward decision-support metrics in tourism governance (OECD, 2024; UN Tourism, 2024). The remainder of the paper outlines data, measures, and empirical strategy, then presents results through distributional evidence, mixed-effects estimates, and probability-scale translations, before discussing implications for threshold-oriented governance and concluding with limitations and future research directions.

2. LITERATURE REVIEW

This chapter establishes the conceptual grounding for the paper's core logic. Objective accessibility is treated as a structural constraint that can enter service evaluation through arrival friction; very low ratings are interpreted as diagnostic and punitive signals rather than marginal dissatisfaction, and travel purpose is expected to shape evaluative sensitivity in ways that matter for an implementable threshold governance rule. The review therefore connects mobility context, online evaluation behaviour, and indicator translation, because the paper's contribution depends not only on estimating an association but also on making it decision-ready under capacity constraints.

2.1. Accessibility as a Structural Constraint in Service Evaluation

Hospitality experiences are embedded in a mobility system. Even when the on-property encounter is stable, travellers incur time costs, transfer burden, and last-mile uncertainty, and these frictions can become part of what is retrospectively evaluated as the "stay". In tourism research, accessibility is therefore not merely a descriptive attribute but a structural condition that shapes feasible choices and reallocates demand when connectivity changes. Evidence from a natural experiment in Xiamen shows that improved intracity public transportation increased demand for nearby hotels, illustrating that accessibility can move accommodation outcomes through a friction-reduction channel (Zheng et al., 2022). Complementary destination-level evidence links transport improvements to tourism redistribution and growth, reinforcing the view that accessibility can operate as a binding constraint rather than a background correlate (Boto-García, 2023; Wardana, Suryantini, & Putra, 2025). This interpretation is consistent with work that defines destination accessibility as a generalised travel-cost concept anchored in travel time from access gates such as airports and rail stations, because the burden of reaching a place can shape visitor flows and experience quality before any in-destination service encounter occurs (Coppola, Borruso, & Murgante, 2020).

Under this framing, minutes and reliability are analytically meaningful because they represent the friction travellers must absorb to convert arrival into an evaluable stay. However, review-based studies often measure "location" using subjective location scores or coarse spatial units, which can conflate structural access constraints with amenity preferences and neighbourhood affect, and therefore translate poorly into implementable routines. This measurement ambiguity motivates an objective minutes-based indicator that can be expressed in operational ranges for triage, and a minimal node taxonomy, such as airport versus urban transit, provides an implementable way to anchor arrival friction to concrete mobility gateways.

Spatially enabled analyses of hotel reviews provide further support for moving from subjective "location" toward measurable constraints. Research integrating geographical factors with review-derived outcomes shows that accessibility-related features can be associated with satisfaction signals, suggesting that location context enters evaluation through measurable frictions rather than neighbourhood sentiment alone (Kashyap & Hong, 2025). Taken together, the literature supports conceptualising accessibility as a structural constraint that enters evaluation through arrival friction and time reliability, which justifies operationalising accessibility as minutes to the nearest transport node rather than as a purely perceptual location score.

2.2. Low Ratings as Diagnostic Signals and Penalty Logic

The paper focuses on very low ratings because online review environments are characterised by asymmetric information processing. A robust mechanism is negativity bias, where negative information is perceived as more diagnostic and can exert disproportionate influence on judgment (Qahri-Saremi & Montazemi, 2023). In platform markets, such asymmetry matters because ratings operate as reputation signals that feed into choice and shape demand. In hotel contexts, experimental evidence indicates that negative ratings and reviews affect booking consideration, supporting the relevance of low-end outcomes for reputational risk even when they are not the modal rating (Gavilan et al., 2018). More broadly, hospitality and tourism research treats eWOM as a consequential informational channel that influences expectations, perceived risk, and decision-making, rather than as a purely expressive post-consumption record (Cantallops & Salvi, 2014; Litvin, Goldsmith, & Pan, 2008).

Classic platform evidence further supports a penalty interpretation of low-end ratings by showing that consumer responses to ratings are often nonlinear and threshold sensitive. One-star reviews can have larger marginal effects than five-star reviews in a prominent eWOM setting, which is consistent with asymmetric diagnosticity (Chevalier & Mayzlin, 2006), while field evidence indicates that consumers respond sharply to boundary-crossing signals rather than treating rating scales as linear (Anderson & Magruder, 2012). In hospitality markets, evidence from China-based travel platforms links online user reviews to hotel room sales, reinforcing that review signals map onto market outcomes (Ye, Law, & Gu, 2009). Additional hotel-focused studies also show that online reviews can shift consumers' consideration sets and perceptions even before direct experience, which is consistent with the idea that negative signals function as risk flags during search (Sparks & Browning, 2011; Vermeulen & Seegers, 2009). Reviews of the hospitality eWOM literature similarly argue that managerial relevance often lies in understanding negative signals and their consequences, not only average sentiment (Cantallops & Salvi, 2014; Schuckert, Liu, & Law, 2015).

These strands motivate a penalty logic for very low ratings. Low-end ratings reflect evaluations where travellers interpret an experience as unacceptable rather than merely imperfect, and such outcomes are more likely to trigger avoidance and complaint diffusion. For governance, the low-rating tail is therefore an appropriate screening target because it concentrates high-salience failures and can be translated into monitoring priorities under limited capacity.

2.3. Travel-Purpose Heterogeneity as an Evaluative Mechanism

Travel purpose is expected to moderate the accessibility–low-rating relationship because business and leisure travellers differ in constraints, scheduling flexibility, and the salience of time reliability. Empirically, evidence from hotel reviews indicates that leisure travellers tend to provide higher ratings and more positive emotional tone than business travellers, while business ratings can be more dispersed, implying different evaluative processes across segments (Kim et al., 2023). Recent work focusing directly on trip type similarly documents systematic differences between business and leisure ratings and links these differences to work-related concerns and evaluative standards (Schiessl et al., 2026). These findings justify modelling travel purpose as a theoretically grounded moderator rather than an exploratory split.

The mechanism is especially relevant for accessibility. Arrival friction is evaluated relative to expectations and constraints, and work-constrained itineraries make minutes and reliability more binding because delays can cascade into schedule disruption, transaction friction, and loss of control. Under this logic, the same transfer uncertainty is more likely to be interpreted as blameworthy failure by business travellers than by leisure travellers, which supports expecting segment-dependent penalties rather than a homogeneous accessibility effect. This heterogeneity is also consequential for threshold governance, because it implies that the same monitoring trigger can warrant different escalation intensity as market mix shifts. At the same time, travel-purpose labels in platform data are proxies and may be missing or inconsistently coded, which reinforces the value of probability-scale interpretation and uncertainty transparency when the objective is prioritisation rather than exhaustive segmentation.

2.4. Synthesis and Implications for the Empirical Design

Taken together, the literature supports a coherent conceptual expectation. Accessibility can be treated as a structural constraint that enters evaluation through arrival friction and time reliability, which justifies measuring it in minutes rather than relying solely on subjective location scores (Kashyap & Hong, 2025; Zheng et al., 2022). Low ratings can be treated as diagnostic penalty signals in platform environments, which justifies modelling the low-rating tail as a governance-relevant risk outcome (Chevalier & Mayzlin, 2006; Qahri-Saremi & Montazemi, 2023). Travel purpose provides a defensible moderator because business and leisure travellers differ in evaluative standards and sensitivity to time and reliability constraints, implying segment-dependent penalties (Kim et al., 2023; Schiessl et al., 2026).

The remaining step is translation. Reviews of the hospitality eWOM literature repeatedly note that managerial value depends on turning review signals into actionable decision support rather than treating them as descriptive text alone (Cantallops & Salvi, 2014; Schuckert et al., 2015). Contemporary tourism governance frameworks likewise emphasise decision-support indicators that can be enacted as routines rather than diffuse reporting (OECD, 2024; UN Tourism, 2024). This orientation motivates an empirical strategy that combines threshold comparisons, which map naturally onto triage, with probability-scale translations that express model outputs as interpretable risk differences. In this sense, the minutes-based indicator is designed not only to strengthen measurement validity but also to enable an indicator-style governance logic in which continuous signals are translated into monitoring triggers under capacity constraints. Accordingly, the next chapter details the data, variable construction, and modelling strategy used to implement this translation in a review-based setting.

3. DATA, MEASURES, AND EMPIRICAL STRATEGY

3.1. Data Sources, Linkage, and Analytic Structure

This study integrates two Ctrip datasets for hotels located in Malaysia, covering reviews posted between October 2022 and December 2025. The first dataset is a review-level file containing ratings and review metadata, and the second dataset is a hotel-level file containing basic hotel attributes and a semi-structured proximity descriptor field. The two files are linked through a harmonised hotel identifier, matching the review-side hotel ID to the hotel-side hotel ID. The unit of analysis is the individual review, while the data structure is hierarchical because multiple reviews are associated with the same hotel. Since within-hotel dependence can lead to overstated precision if ignored, the empirical models use mixed-effects logistic regression with a hotel random intercept, which is a standard approach for clustered observations (Bates, Mächler, Bolker, & Walker, 2015; Meteyard & Davies, 2020).

After standardisation and removal of records with missing hotel identifiers or missing ratings, the analysis further restricts the sample to reviews whose hotels have parsable accessibility information from the proximity descriptor field, which yields a model-ready dataset with non-missing nearest-node minutes and node type. The resulting analytic sample contains 346,988 reviews from 792 hotels. Table 1 summarises the analytic sample and the distributions of all variables used in the models. At the review level, the mean rating is 4.34 (SD = 0.87) and the share of low ratings, defined as ratings below 3, is 6.59%. At the hotel level, the mean nearest-node travel time is 15.48 minutes (SD = 9.65), the median is 13 minutes, and the interquartile range is 9 to 21 minutes, while 40.03% of hotels are classified as airport-nearest and 59.97% as transit-nearest. These descriptive statistics are reported to make the empirical scope and modelling inputs transparent, consistent with recommendations that emphasise clear reporting of random-effects structures and data hierarchy (Meteyard & Davies, 2020).

Table 1. Variables and descriptive statistics.

Panel A. Review-level descriptives	
Statistic	**Value**
Number of reviews	346,988
Mean rating (SD)	4.34 (0.87)
Low rating (rating < 3)	6.59%
Mean review length (Characters)	130.45
Business share	16.83%
Number of hotels	792

Panel B. Hotel-level descriptives	
Statistic	**Value**
Number of hotels	792
Nearest node type: Airport share	40.03%
Nearest node type: Transit share	59.97%
Mean nearest-node minutes (SD)	15.48 (9.65)
Minutes distribution (P25; median; P75)	9; 13; 21
Mean star rating	3.64
Mean listed price	464.63
Mean hotel review count	653.88

Note: Entries report descriptive statistics for the analytic sample. LowRating is defined as a rating < 3. Nearest-node minutes are derived from platform proximity descriptors and defined as the minimum travel time to the nearest airport- or transit-related node for each hotel.

3.2. Constructing Objective Accessibility from Proximity Descriptors

Objective accessibility is operationalised as an estimated travel time in minutes from each hotel to its nearest transport node as described in the hotel-level proximity field. This operationalisation is consistent with evidence that transport accessibility can be an actionable condition within tourism systems and may translate into measurable shifts in accommodation-related market outcomes (Zheng et al., 2022). Accessibility is derived from the semi-structured proximity descriptor using a rule-based, auditable procedure. Proximity strings are normalised and split into items using common delimiters, after which each item is classified into either airport-related or transit-related nodes via keyword rules. Minutes are extracted from travel-time expressions using English patterns such as "X min" and "X minutes" (including common variants and spacing). For each hotel and node type, the nearest time is defined as the minimum minutes across items within that type. The hotel's nearest-node accessibility is subsequently defined as the minimum minutes across airport and transit, and the corresponding node type is recorded as an attribute for stratification. Accessibility is constructed at the hotel level and merged to the review level through the hotel identifier so that each review inherits its hotel's accessibility attributes. To support both flexible modelling and threshold-oriented interpretation, the analysis uses a continuous minute's measure as well as a discretised four-band measure (0–10, 10–20, 20–40, and >40 minutes). Distributional visualisations are reported in the Results section; for the density plot, minutes are restricted to 0–40 solely to improve readability of the displayed distribution, while the modelling covariates retain the full minutes range unless a specification explicitly imposes a restriction.

3.3. Measures

The dependent variable (LowRating) is a binary indicator equal to 1 when a review's rating is below 3 and 0 otherwise. This strict threshold is motivated by research showing that negative online review information can be disproportionately diagnostic in digital decision contexts, which supports modelling the punitive tail explicitly rather than relying solely on mean shifts (Qahri-Saremi & Montazemi, 2023).

Travel-purpose heterogeneity is captured through a business indicator constructed from the travel-type field when available. The indicator is coded as 1 when the travel-type label contains business-related tokens and as 0 otherwise. When travel-type information is unavailable in the raw file, the indicator is coded as 0, and it is treated as a conservative proxy rather than a complete taxonomy. This design is grounded in evidence that business and leisure travellers can differ systematically in their online evaluations of hotel services, which makes trip purpose a theoretically defensible moderator (Kim et al., 2023).

The models also include review- and hotel-level controls. Review length is measured as the character count of the review text and is entered as $\log(1 + review\_length)$ to reduce skew and capture reporting intensity. Hotel-level controls include star rating, listed price, and hotel review count, where price and review count enter as $\log(1 + price)$ and $\log(1 + review\_count)$ to stabilise scale and limit leverage from extreme values.

3.4. Empirical Strategy and Model Equations

The empirical analysis uses mixed-effects logistic regression to reflect the nested structure of reviews within hotels. Let (i) index reviews and (h) index hotels. The baseline threshold specification is:

$$\text{logit}(\Pr(LowRating_{ih} = 1)) = \beta_0 + \sum_k \beta_k \ 1(MinutesBand_h = k) + X_{ih}\gamma + u_h, \text{ with } u_h \sim \mathcal{N}(0, \sigma_u^2) \quad (1)$$

Estimation is conducted in R using maximum likelihood for mixed-effects logistic regression, and models are fitted with the bobyqa optimizer with an increased iteration budget to support stable convergence in specifications that include interactions and spline terms (Bates et al., 2015).

Where MinutesBand is defined at the hotel level with 0 to 10 minutes as the reference category, (\mathbf{X}_{ih}) contains the controls described above, and (u_h) is a hotel random intercept. Estimation is implemented in lme4 and reported as odds ratios with confidence intervals to facilitate comparison across bands (Bates et al., 2015).

To assess heterogeneity by travel purpose, the model incorporates an interaction between MinutesBand and the business indicator.

$$\text{logit}(\Pr(LowRating_{ih} = 1)) = \beta_0 + \sum_k \beta_k \ 1(MinutesBand_h = k) + \delta \, Business_{ih} + \sum_k \theta_k \ 1(MinutesBand_h = k) \, Business_{ih} + X_{ih}\gamma + u_h \quad (2)$$

Because band-based coefficients may not fully capture nonlinear patterns in accessibility, a flexible specification is also estimated using a natural spline of minutes and an interaction with business status. Minutes are winsorised at 55 minutes to reduce sensitivity to extreme tail values.

$$\text{logit}(\Pr(LowRating_{ih} = 1)) = \beta_0 + \text{ns!}\left(minutes_h^{cap}, df = 4\right)\beta + Business_{ih} \cdot \text{ns!}\left(minutes_h^{cap}, df = 4\right)\theta + X_{ih}\gamma + u_h, \text{ where } minutes_h^{cap} = \min(minutes_h, 55) \quad (3)$$

Predicted probabilities are computed from the fitted linear predictor via $\hat{p} = (1 + \exp(-\hat{\eta}))^{-1}$., holding controls at their median values. Uncertainty bands for predicted probabilities are derived from the fixed-effects variance–covariance matrix using a Wald approximation, which provides a transparent model-based summary rather than a full simulation over the random-effects distribution. This probability-based reporting convention is consistent with contemporary guidance on interpreting nonlinear and interaction models via predicted quantities (Arel-Bundock et al., 2024).

3.5. Reporting and Reproducibility Conventions

All data processing, modelling, and visualisation are implemented in R (R Core Team, 2025). Data cleaning and variable construction rely on the tidyverse ecosystem for consistent data manipulation and string processing (Wickham et al., 2019), while visualisations are produced using ggplot2's layered grammar (Wickham, 2016). Mixed-effects logistic regressions are estimated using lme4 (Bates et al., 2015), and model-based interpretation is supported through predicted quantities in probability space, following established guidance for interpreting nonlinear and interaction models (Arel-Bundock et al., 2024). High-resolution figure exports are generated using the Ragg graphics devices to ensure consistent raster output suitable for journal production workflows (Pedersen & Shemanarev, 2025). When predicted probabilities are plotted for nonlinear specifications, uncertainty bands are derived from the fixed-effects variance–covariance matrix using a Wald approximation, which provides a transparent model-based summary of estimation uncertainty rather than a full simulation over the random-effects distribution (Meteyard & Davies, 2020). In addition, bin-level descriptive rates are accompanied by Wilson score intervals to improve coverage relative to normal-approximation intervals for proportions (Wilson, 1927).

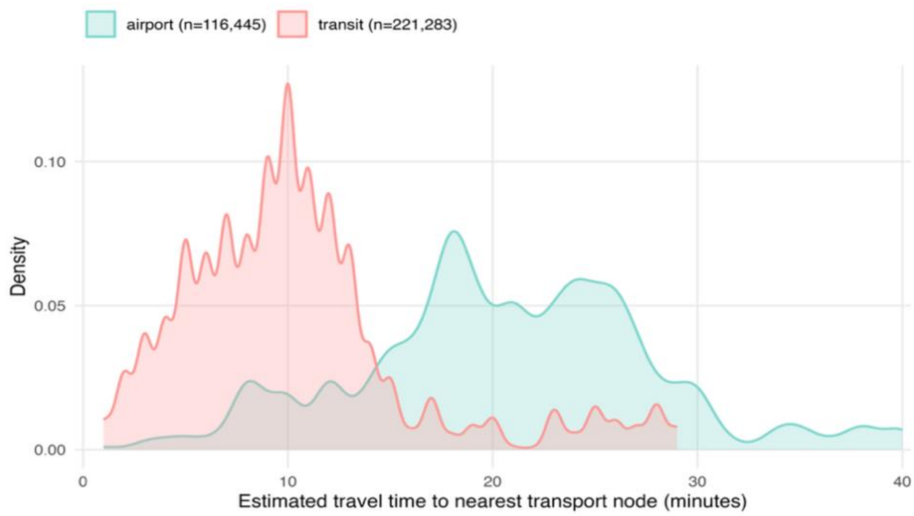

Figure 2. Accessibility distribution by nearest transport node type.

Note: Kernel density of nearest-node travel time (Minutes). X-axis limited to 0–40 for readability.

4. RESULTS

4.1. Descriptive Patterns

Figure 2 summarises the distribution of estimated travel time to the nearest transport node, distinguishing whether the nearest node is classified as an airport or transit. Because the density plot is rendered on the 0-to-40-minute window for readability, the legend counts correspond to observations within that plotted range rather than the full minutes of support. Within this displayed window, the transit-nearest distribution concentrates at shorter travel times, whereas the airport-nearest distribution is shifted toward longer travel times and spreads more broadly. This descriptive contrast matters because it clarifies that objective accessibility is heterogeneous across hotels, and it is consistent with recent evidence that spatial context and proximity-related factors can shape review-derived signals of satisfaction rather than functioning as neutral background conditions (Kashyap & Hong, 2025). It also aligns with accommodation market evidence showing that improvements in intracity public transportation can shift hotel demand by changing accessibility, which supports treating minutes as a structurally meaningful variable rather than a cosmetic descriptor (Zheng et al., 2022). Importantly, the 0-to-40-minute restriction is applied only to the plotted density for readability and does not alter the estimation sample used in the regression models.

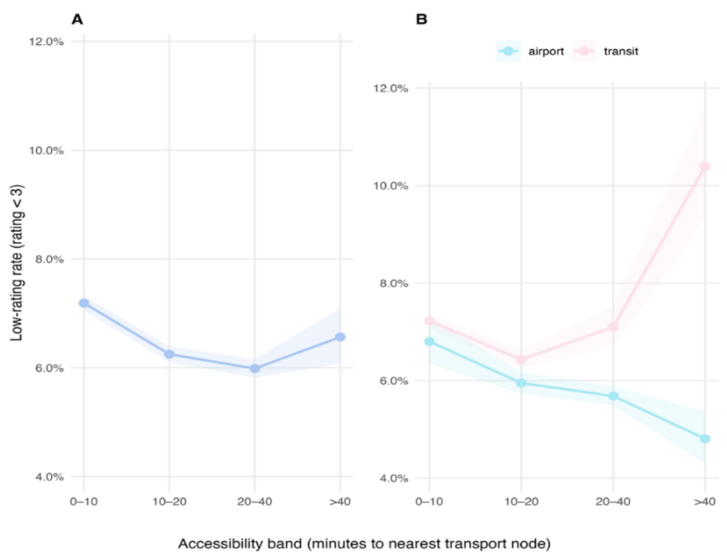

Figure 3. Low rating across accessibility bands.

Note: Panel A reports the low-rating rate (rating < 3) across four accessibility bands based on minutes to the nearest transport node (0–10, 10–20, 20–40, and >40). Panel B reports the same band-level rates stratified by nearest node type (airport vs transit). Shaded areas denote 95% Wilson confidence intervals.

Figure 3 then reports low-rating rates across four accessibility bands, defined as 0 to 10, 10 to 20, 20 to 40, and above 40 minutes. Panel A presents the overall band-level pattern, whereas Panel B reports the same band structure separately for airport-nearest and transit-nearest hotels. The figure is designed to provide a descriptive bridge between the continuous minute distribution in Figure 2 and the threshold-based modelling strategy. Uncertainty around each band-level rate is displayed using Wilson score intervals, which are widely cited for providing more reliable coverage than normal-approximation intervals for binomial proportions (Wilson, 1927). In this way, Figure 3 establishes that the band logic is empirically interpretable before turning to model-based estimates.

4.2. Main Regression Results

Table 2 reports mixed-effects logistic regressions predicting whether a review is a low rating, defined as a rating below 3. The models include a hotel random intercept to reflect the nested structure of reviews within hotels, and results are reported as odds ratios with confidence intervals, consistent with standard practice for generalised mixed-effects models and recommended reporting guidance (Bates et al., 2015; Meteyard & Davies, 2020). Model fit improves as controls are introduced, showing a decrease in AIC from 161,039 in Model 1 to 156,675 in Model 3.

Table 2. Mixed-effects logistic regression predicting low rating (Rating < 3).

Term	M1 OR [95% CI]	M1 p	M2 OR [95% CI]	M2 p	M3 OR [95% CI]	M3 p
10–20 min vs 0–10	0.81 [0.72, 0.91]	< 0.001	0.80 [0.71, 0.90]	< 0.001	0.80 [0.71, 0.89]	< 0.001
20–40 min vs 0–10	0.75 [0.65, 0.85]	< 0.001	0.73 [0.64, 0.84]	< 0.001	0.73 [0.64, 0.82]	< 0.001
>40 min vs 0–10	0.75 [0.52, 1.07]	0.114	0.69 [0.47, 1.00]	0.048	0.98 [0.71, 1.37]	0.920
Business (1 = yes)			1.22 [1.18, 1.27]	< 0.001	1.22 [1.17, 1.27]	< 0.001
Review length (log1p)			1.51 [1.49, 1.53]	< 0.001	1.51 [1.49, 1.53]	< 0.001
Star rating					0.99 [0.92, 1.06]	0.787
Price (log1p)					0.61 [0.55, 0.68]	< 0.001
Hotel review count (log1p)					0.94 [0.88, 1.01]	0.082

Note: Entries are odds ratios with 95% confidence intervals in brackets. The reference category for accessibility bands is 0–10 minutes. All models are mixed-effects logistic regressions with a hotel random intercept. M1 includes accessibility bands only. M2 adds business and review length. M3 further adds hotel attributes (star rating, log price, log hotel review count).

In the fully adjusted specification, accessibility bands exhibit a structured pattern relative to the 0-to-10-minute reference group. The 10-to-20-minute band is associated with lower odds of a low rating, with an odds ratio of 0.80 and a 95% confidence interval from 0.71 to 0.89 ($p<.001$). The 20-to-40-minute band shows a similar association, with an odds ratio of 0.73 and a 95% confidence interval from 0.64 to 0.82 ($p<.001$). By contrast, the above 40-minute band is close to unity in the fully adjusted model, with an odds ratio of 0.98 and a 95% confidence interval from 0.71 to 1.37 ($p=.920$). The covariates in Model 3 indicate that business-related stays have higher odds of a low rating than non-business stays, with an odds ratio of 1.22 ($p<.001$), while review length is positively associated with low-rating odds, with an odds ratio of 1.51 ($p<.001$). The association between review verbosity and rating outcomes is consistent with recent evidence showing that review text length relates systematically to numerical scores and that the relationship can involve nonlinearities, which supports controlling for verbosity when modelling rating extremes (Mellinas & Leoni, 2025).

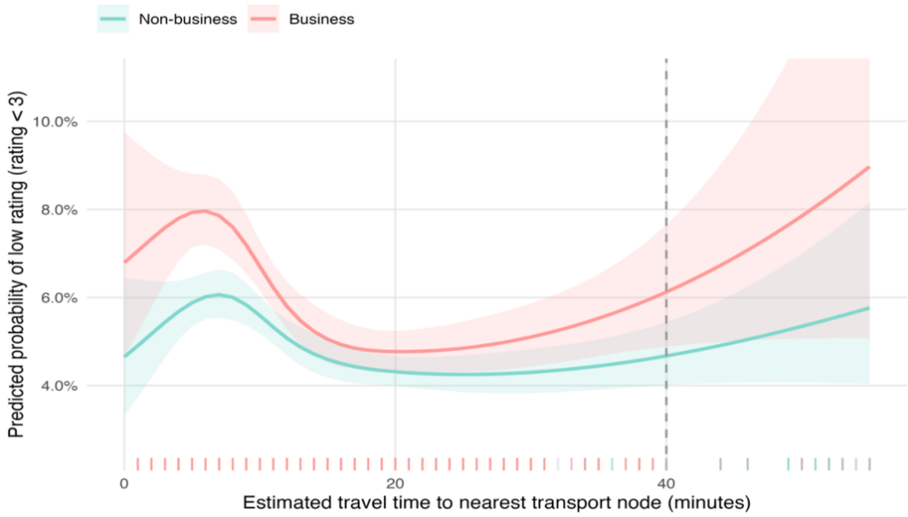

Figure 4. Predicted low-risk rating by accessibility minutes and travel purpose.

Note: Predicted probabilities (rating < 3) are derived from a mixed-effects logistic model with a hotel random intercept, modelling accessibility as a spline of travel time capped at 55 minutes. Shaded bands indicate 95% confidence intervals. The dashed vertical line marks 40 minutes.

Figure 4 translates the regression specification into predicted probabilities across the minutes scale, using a flexible nonlinear model that permits the association between minutes and low-rating risk to vary across the range. Reporting predicted probabilities is particularly useful here because probabilities are easier to interpret than odds ratios when nonlinearity and group differences are present, and this strategy aligns with contemporary guidance to interpret statistical models using predicted quantities and contrasts (Arel-Bundock et al., 2024). Consistent with the threshold specification in Table 2, Figure 4 provides a complementary probability-scale summary that allows the minutes association to vary flexibly across the support, with covariates held at their median values. At the 40-minute reference point, the predicted probability of a low rating is 6.13% (95% CI: 4.89%–7.65%) for business stays and 4.67% (95% CI: 4.01%–5.44%) for non-business stays. By 55 minutes, the corresponding probabilities are 8.98% (95% CI: 5.08%–15.38%) and 5.76% (95% CI: 4.04%–8.16%), implying an absolute business–non-business gap that widens from 1.45 to 3.21 percentage points. These probability translations are presented alongside uncertainty bands and rug marks to keep interpretation anchored in the precision of model-based estimates, which is particularly salient in online review contexts where negative information can carry disproportionate diagnostic weight for consumer judgement (Qahri-Saremi & Montazemi, 2023). More broadly, experimental evidence also suggests that bad ratings can exert a strong influence in early-stage hotel choice, reinforcing why probability-scale differences for low-rating risk are meaningful to report (Gavilan et al., 2018).

Table 3. Interaction model: accessibility bands × business travel.

Term	OR [95% CI]	p
10–20 min vs 0–10 (Non-business)	0.82 [0.73, 0.91]	< 0.001
20–40 min vs 0–10 (Non-business)	0.74 [0.65, 0.84]	< 0.001
>40 min vs 0–10 (Non-business)	0.98 [0.70, 1.37]	0.908
Business vs non-business (at 0–10)	1.30 [1.23, 1.38]	< 0.001
Review length (log1p)	1.51 [1.49, 1.53]	< 0.001
Star rating	0.99 [0.92, 1.07]	0.802
Price (log1p)	0.61 [0.55, 0.67]	< 0.001
Hotel review count (log1p)	0.94 [0.88, 1.01]	0.081
Interaction: (10–20) × Business	0.86 [0.79, 0.94]	< 0.001
Interaction: (20–40) × Business	0.90 [0.81, 1.00]	0.054
Interaction: (>40) × Business	1.27 [0.82, 1.99]	0.286

Note: Entries are odds ratios with 95% confidence intervals in brackets. Minutes-band main effects are interpreted for the non-business group (Business = 0), with 0–10 minutes as the reference band. The "Business vs non-business (at 0–10)" term compares travel purposes within the 0–10-minute band. Interaction terms indicate how minutes-band effects differ for business travellers. The model includes a hotel random intercept.

4.3. Heterogeneity and Threshold Governance Emphasis

Table 3 extends the main specification by interacting accessibility bands with travel purpose, operationalised as business versus non-business stays. This heterogeneity test is aligned with evidence that business and leisure travellers can systematically differ in hotel rating behaviour and evaluative standards, which makes travel purpose a theoretically grounded moderator rather than an ad hoc split (Kim et al., 2023; Schiessl et al., 2026). In Table 3, the main effects for accessibility bands describe differences for the non-business group relative to the 0-to-10-minute baseline. For non-business stays, the odds ratio is 0.82 ($p<0.001$) for 10 to 20 minutes and 0.74 ($p<0.001$) for 20 to 40 minutes, while the above 40-minute band is 0.98 ($p=0.908$). The main effect of the business at the baseline band is 1.30 ($p<0.001$), indicating higher baseline low-rating odds for business stays at 0 to 10 minutes. The interaction terms indicate how the band differentials shift for business stays, with an odds ratio of 0.86 ($p<0.001$) for the 10 to 20 interaction, 0.90 ($p=0.054$) for the 20 to 40 interaction, and 1.27 ($p=0.286$) for the above 40 interaction, with wider uncertainty for the farthest band.

A within-group translation for business stays can be obtained by combining the relevant main and interaction terms. Under that conditional translation, the estimated odds ratio for business stays is approximately 0.70 for 10 to 20 minutes relative to 0 to 10, with a 95% confidence interval from 0.62 to 0.80, and approximately 0.67 for 20 to 40 minutes relative to 0 to 10, with a 95% confidence interval from 0.58 to 0.78. For the above 40-minute band, the conditional odds ratio for business stays is approximately 1.25 with a 95% confidence interval from 0.73 to 2.13, indicating a point estimate above unity alongside substantial uncertainty.

Figure 4 complements this interaction evidence by expressing heterogeneity in probability space. Because the figure plots predicted low-rating probabilities separately for business and non-business stays, the separation between curves can be interpreted as an absolute risk gap rather than as an interaction coefficient that requires additional transformation, and this mode of communication follows recommendations to foreground predicted quantities when interpreting nonlinear and moderated relationships (Arel-Bundock et al., 2024). In probability space, the same heterogeneity is visible at the policy-relevant cut-points, where the business–non-business risk gap is 1.45 percentage points at 40 minutes and widens to 3.21 percentage points by 55 minutes, while uncertainty expands in the far-right tail. The marked emphasis on the 40-minute cut-point connects directly to the paper's threshold governance framing because it defines an operational boundary at which monitoring intensity and mitigation prioritisation can be escalated under capacity constraints, while still preserving uncertainty-aware judgement through confidence bands. This indicator-oriented logic aligns with policy-facing work that argues tourism governance benefits from decision-ready indicators and clearly interpretable thresholds, as reflected in OECD's emphasis on strengthening the evidence base and indicator toolkits for destination-level decision making (OECD, 2024) and UN Tourism's Statistical Framework for Measuring the Sustainability of Tourism, which explicitly positions comparable statistical structures as decision-support infrastructure for governance (UN Tourism, 2024).

To maintain a clean separation between results and interpretation, the next chapter develops the underlying mechanisms and discusses how the threshold logic can be translated into implementable monitoring and response routines.

5. DISCUSSION AND CONCLUSION

5.1. Reconnecting The Findings to the Research Questions and the Policy Window

This study asked whether objective accessibility operates as a structural constraint that is penalised in the low-rating tail, whether the penalty differs by travel purpose, and whether the pattern supports an actionable threshold rule. Read as a single chain, the results are internally consistent. The minutes indicator varies meaningfully across hotels and functions as a tractable proxy for arrival friction, which fits evidence that mobility conditions shape tourism and accommodation outcomes (Kashyap & Hong, 2025; Zheng et al., 2022). The band and spline representations further justify a threshold lens because the association is not well described by a single linear slope. Crucially, the probability translation makes the governance meaning legible by expressing effects as decision-relevant risk differences rather than coefficient changes, which is increasingly recommended when nonlinearity and moderation matter (Arel-Bundock et al., 2024).

This translation is especially relevant under capacity constraints, where a useful indicator must be enactable as routine monitoring. In that sense, the 40-minute cut-point can be treated as an operational trigger rather than a deterministic boundary. The predicted probability gap between business and non-business stays is 1.45 percentage points at 40 minutes and widens to 3.21 percentage points by 55 minutes, while uncertainty expands in the far-right tail where empirical support is thinner. Framed as a trigger, the threshold supports prioritisation without overstating precision, which aligns with policy-facing guidance that emphasises comparable measures and implementable governance routines (OECD, 2024; UN Tourism, 2024).

5.2. Why Low Ratings are a Governance Target

Focusing on the low-rating tail is defensible because online review environments are shaped by information asymmetry and selective attention. Negative signals can be disproportionately diagnostic in judgment, which makes changes in the probability of very low ratings practically meaningful even when low ratings are infrequent (Qahri-Saremi & Montazemi, 2023). In hotel choice settings, ratings and review signals shape evaluation and consideration, so the low end of the distribution is a reasonable proxy for reputational risk (Gavilan et al., 2018). The paper's aim is therefore not to claim strong causality, but to show that an objectively measurable accessibility signal can be used as a parsimonious and auditable risk screen.

5.3. Business Travel as a Sensitivity Amplifier

Travel purpose moderates the accessibility signal. Businesses exhibit higher baseline low-rating risk, and the gap widens as minutes increase into higher-friction ranges. This pattern is consistent with evidence that business and leisure travellers differ in evaluative standards and rating behaviour in online hotel reviews (Kim et al., 2023; Schiessl et al., 2026). It is also consistent with a time sensitivity logic in which reliability and predictability become binding constraints for work-constrained travel, increasing the likelihood that arrival friction is translated into punitive evaluation. Even if travel purpose remains an imperfect proxy, it offers a governance-relevant calibration because it indicates that the same structural friction can attract different penalties as market mix shifts.

5.4. Contributions

The study contributes in three ways that match a policy window, plus an implementable indicator agenda. Methodologically, it shows how platform proximity descriptors can be converted into an objective minutes-to-node indicator using auditable parsing rules, reducing reliance on subjective location ratings and enabling replication without heavy geospatial infrastructure. Empirically, it documents a structured accessibility pattern in the low-rating tail and shows that the pattern differs by travel purpose, linking a structural constraint to penalty logic and segment sensitivity within one design. Translationally, it reframes the association as threshold governance by reporting probability-scale risk and anchoring a monitoring trigger around an interpretable cut-point, which aligns with decision-support approaches in contemporary tourism governance (OECD, 2024; UN Tourism, 2024).

5.5. Policy and Managerial Implications Through a Threshold Lens

The threshold framing implies a prioritisation rule rather than a generic claim that location matters. At the destination level, the 40-minute trigger can be used to flag higher-friction zones and coordinate targeted measures that reduce arrival uncertainty and transfer burden, consistent with viewing accessibility as an actionable system lever (Zheng et al., 2022). Because responsibility is distributed across hotels, transport operators, and information intermediaries, a shared indicator can lower coordination costs by creating a common language for triage and sequencing. At the hotel level, the implication centres on expectation management and reliability. When a property falls into higher-friction ranges, clearer arrival instructions, more predictable transfer options, and standardised guidance become more consequential, particularly for business-heavy demand. These interventions do not change distance, yet they can reduce uncertainty and the likelihood that friction is reinterpreted as service failure. The threshold logic also supports internal allocation decisions by indicating where investment in pre-arrival communication and contingency handling is most likely to protect against low-tail reputational shocks.

At the platform level, standardising proximity information and making arrival guidance more consistent can reduce navigational friction at the point of planning and arrival. Since cue diagnosticity implies that contextual cues interact with reviews to shape evaluation, platform-side information design is not neutral in how experiences are framed and later rated (Qahri-Saremi & Montazemi, 2023). In practice, structured proximity fields and searchable arrival guidance can complement hotel-side efforts by reducing ambiguity.

5.6. Limitations and Future Research

Several limits bound interpretation. The analysis uses one platform and a rule-based minutes extraction, so replication across platforms and validation against GIS-based travel-time estimates would strengthen external validity. The results are associational, so quasi-experimental extensions using transport shocks or policy changes are a natural next step, building on work that treats transport changes as plausibly exogenous shifts in hotel markets (Zheng et al., 2022). Travel purpose may also be missing or misclassified for some observations, which motivates robustness checks under alternative coding and purpose inference strategies. Future studies could further connect threshold-based low-rating risk to behavioural outcomes such as conversion or revenue effects, extending the governance logic from risk screening to performance consequences (Gavilan et al., 2018).

5.7. Conclusion

This study translates platform-derived accessibility information into a decision-ready threshold rule for low-risk rating in Malaysian hotels. By constructing an objective minutes-to-transport indicator from proximity descriptors and translating estimates into probability-scale risk differences, it provides an interpretable monitoring trigger around a 40-minute cut-point and shows that risk differences are amplified for business stays. More broadly, it illustrates how review analytics can support governance not only through text interpretation, but also through auditable indicators that can be enacted as routines under capacity constraints, in line with tourism governance frameworks that emphasise comparable, decision-support measures (OECD, 2024; UN Tourism, 2024).

Funding: This study received no funding.
Institutional Review Board Statement: The study used secondary data derived from publicly available online reviews on Ctrip (Trip.com). All analyses were conducted on anonymized and aggregated data, and no identifiable personal information was collected or reported. As the study involved secondary analysis of public, de-identified information with no direct human participation, ethical review and approval were not required.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Anderson, M., & Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *The Economic Journal*, *122*(563), 957-989. https://doi.org/10.1111/j.1468-0297.2012.02512.x

Arel-Bundock, V., Greifer, N., & Heiss, A. (2024). How to interpret statistical models using marginaleffects for R and Python. *Journal of Statistical Software*, *111*(9), 1-32. https://doi.org/10.18637/jss.v111.i09

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boto-García, D. (2023). High-speed rail and tourism: Evidence from Spain. *Tourism Economics*, *29*(2), 381–401.

Cantallops, A. S., & Salvi, F. (2014). New consumer behavior: A review of research on eWOM and hotels. *International Journal of Hospitality Management*, *36*, 41-51. https://doi.org/10.1016/j.ijhm.2013.08.007

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(3), 345-354. https://doi.org/10.1509/jmkr.43.3.345

Coppola, P., Borruso, G., & Murgante, B. (2020). Accessibility and tourism: A review of the literature. *Sustainability*, *12*(21), 8775.

Ctrip.com. (1999). *Company profile*. Retrieved from https://webresource.ctrip.com/code/engpurehtml/footer/CompanyProfile.html

Gavilan, D., Avello, M., & Martinez-Navarro, G. (2018). The influence of online ratings and reviews on hotel booking consideration. *Tourism Management*, *66*, 53-61. https://doi.org/10.1016/j.tourman.2017.10.018

Kashyap, A., & Hong, S.-Y. (2025). Analysis of hotel reviews and ratings with geographical factors in Seoul: A quantitative approach to understanding tourist satisfaction. *ISPRS International Journal of Geo-Information*, *14*(9), 334. https://doi.org/10.3390/ijgi14090334

Kim, J. M., Ma, H., & Park, S. (2023). Systematic differences in online reviews of hotel services between business and leisure travelers. *Journal of Vacation Marketing*, *29*(2), 189-205. https://doi.org/10.1177/13567667221084373

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, *29*(3), 458-468. https://doi.org/10.1016/j.tourman.2007.05.011

Mellinas, J. P., & Leoni, V. (2025). Beyond words: Unveiling the implications of blank reviews in online rating systems. *Information Technology & Tourism*, *27*(1), 263-284. https://doi.org/10.1007/s40558-024-00300-4

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*, 104092. https://doi.org/10.1016/j.jml.2020.104092

Ministry of Tourism Arts and Culture. (2026). *Official launch of VM2026 in collaboration with the Melaka State Government*. Malaysia: Ministry of Tourism Arts and Culture.

OECD. (2024). *OECD tourism trends and policies 2024*. France: OECD Publishing.

Pedersen, T. L., & Shemanarev, M. (2025). *ragg: Graphic devices based on AGG (R package version 1.5.0)*. USA: Comprehensive R Archive Network.

Qahri-Saremi, H., & Montazemi, A. R. (2023). Negativity bias in the diagnosticity of online review content: The effects of consumers' prior experience and need for cognition. *European Journal of Information Systems*, *32*(4), 717-734. https://doi.org/10.1080/0960085X.2022.2041372

R Core Team. (2025). *R: A language and environment for statistical computing*. Austria: R Foundation for Statistical Computing.

Schiessl, D., Manosso, F. C., Alves, F., & Prado, P. H. M. H. (2026). Hotel ratings: The impact of trip type (business vs leisure). *Tourism Review*, *81*(1), 436-454. https://doi.org/10.1108/TR-11-2023-0807

Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing*, *32*(5), 608-621. https://doi.org/10.1080/10548408.2014.933154

Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, *32*(6), 1310-1323. https://doi.org/10.1016/j.tourman.2010.12.011

Tourism Malaysia. (2026). *VM2026 officially begins: Malaysia rolls out a warm welcome nationwide*. Malaysia: Tourism Malaysia.

Trip.com Group Limited. (2024). *Annual report (Form 20-F) for the fiscal year ended*. https://www1.hkexnews.hk/listedco/listconews/sehk/2025/0411/2025041101009.pdf

UN Tourism. (2024). *Statistical framework for measuring the sustainability of tourism (SF-MST): Final draft (Background document for the 55th session of the UN Statistical Commission)*. New York: United Nations Statistics Division.

Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management*, *30*(1), 123-127. https://doi.org/10.1016/j.tourman.2008.04.008

Wardana, I. M., Suryantini, N. P., & Putra, I. M. (2025). Transportation infrastructure and tourism growth: Evidence from Indonesia. *Tourism Economics, 31*(1), 120–141.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed.). USA: Springer.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., . . . Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association, 22*(158), 209–212. https://doi.org/10.1080/01621459.1927.10502953

Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management, 28*(1), 180-182. https://doi.org/10.1016/j.ijhm.2008.06.011

Zheng, W., Gao, S., & Fang, B. (2022). How does public transportation influence hotel demand? Evidence from a natural experiment. *International Journal of Hospitality Management, 107*, 103338. https://doi.org/10.1016/j.ijhm.2022.103338