Beyond Net Effects: An fsQCA and NCA Reanalysis of Mobile Payment Intention on Trans Jatim Buses

Edwin Pramana[1*], Yosi Kristian[2], Ono Supriadi[3]

[1,2]Informatics Department, Institut Sains dan Teknologi Terpadu Surabaya, Surabaya, East Java, Indonesia
[3]School of Business, Western Sydney University, Surabaya, East Java, Indonesia
E-mail: [1*]epramana@stts.edu, [2]yosi@stts.edu, [3]O.Supriadi@westernsydney.edu.au

(Received: 7 Aug 2026, revised: 30 Aug 2026, accepted: 3 Sep 2026)

Abstract

Mobile payment is now a standard feature of public transport services, yet uptake varies considerably across service contexts. Research on this question has relied mainly on variable-centered methods such as Structural Equation Modeling (SEM), which estimate the net effect of each predictor but do not directly identify equifinal configurations or necessary-condition bottlenecks. This article reports a secondary reanalysis of the respondent-level dataset (n = 432) underlying a previously published extended Unified Theory of Acceptance and Use of Technology (UTAUT) study of Trans Jatim bus passengers. No new respondents were recruited, and the SEM model was not re-estimated here; the previously reported SEM estimates serve only as a net-effect benchmark. That baseline, computed in AMOS 22 on the same data, showed Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), Trust (TR), and Network Externalities (NE) to be significantly and positively associated with Behavioral Intention (BI), while Promotional Activities (PA) were not. The present study adds fuzzy-set Qualitative Comparative Analysis (fsQCA) for sufficient configurations and Necessary Condition Analysis (NCA) for necessary-but-not-sufficient conditions. At the stated calibration and truth-table thresholds, fsQCA returned six empirically supported sufficient configurations for high BI, each combining five of the six conditions. These configurations overlap heavily, and 380 of the 432 cases fall in the all-present truth-table row, so the observed equifinality is bounded by limited empirical diversity. The parsimonious solution is EE + PE, which in Boolean notation denotes two alternative single-condition terms (EE or PE) rather than a joint PE*EE configuration. NCA identified Trust as the condition with the largest observed necessity effect among those tested (CE-FDH d = 0.208, p < 0.001). Promotional Activities were only nominally significant (d = 0.104, p = 0.044) and did not survive Holm or Benjamini-Hochberg adjustment for six simultaneous tests, so that result is reported as tentative. Within this dataset, perceived usefulness and ease of use are the conditions retained under logical minimization, while trust shows the clearest evidence of a bottleneck relationship with behavioral intention.

Keywords: Mobile Payment, Behavioral Intention, Public Transportation, UTAUT, fsQCA, Necessary Condition Analysis.

I. INTRODUCTION

Mobile payment has become part of the everyday fabric of transportation services, promising faster transactions, less dependence on cash, and tighter integration with digital mobility ecosystems [1-3]. In public transport, the stakes are concrete: how passengers pay affects boarding speed, queue length at the door, operator cash-handling costs, and the pace at which a transit system can modernize, and cashless fare policies raise equity questions when they are rolled out unevenly [4-5]. This equity dimension is not analyzed here and is mentioned only as contextual motivation. Availability, however, does not translate into use. Passengers postpone or resist mobile payment when perceived risk, service disruption, or contextual barriers remain salient [6], and a user may regard the system as useful while still hesitating over trust, ease of use, social familiarity, or the limited range of places the payment method is accepted. For operators hoping to expand digital fare collection, this gap between availability and uptake is the practical problem.

Two features of this empirical setting make it well suited to a configurational and necessary-condition reading. First, regional bus systems in emerging economies operate under uneven digital infrastructure, mixed passenger digital literacy, and partial merchant acceptance, so favorable adoption conditions rarely arrive uniformly, and the relevant question is which combinations of conditions are actually present together. Second, such systems still carry a cash-based fallback, which means that any single deficient condition, for example, low trust in transaction reliability, can be enough to keep a

passenger on the incumbent payment method even where every other condition is favorable. Emerging-economy regional transit is therefore an informative site for methods that model conjunctural sufficiency and minimum-level necessity rather than average net effects alone.

An important scope condition should be stated at the outset. The outcome modeled in the source dataset, and therefore in this reanalysis, is self-reported Behavioral Intention to use mobile payment, not observed transaction behavior or actual adoption. Throughout this article, the terms 'intention' and 'behavioral intention' are used deliberately, and conclusions are restricted to intention unless actual usage is explicitly under discussion.

Technology adoption research has typically approached such questions with variable-centered models. The Unified Theory of Acceptance and Use of Technology (UTAUT) is used here not because it is widely cited but because its core constructs map directly onto the decision a bus passenger faces at the point of boarding: whether the payment method delivers a practical benefit (Performance Expectancy), whether it can be operated quickly under time pressure in a queue (Effort Expectancy), and whether its use is socially visible and endorsed in a shared, observable setting (Social Influence) [7-8]. The facilitating-conditions construct of the base model is less informative in this setting, because the enabling infrastructure is supplied by the operator rather than chosen by the passenger, and it was therefore not carried into the source model.

Consumer and mobile service applications have extended the base model with contextual factors including Trust, Perceived Security, Network Externalities, Promotional Activities, Perceived Risk, and Financial Incentives [8-14]. The three extensions retained in the source model were selected on substantive grounds. Trust addresses the specific vulnerability of paying before boarding a vehicle, where a failed or unconfirmed transaction is immediately consequential [10], [15]. Network Externalities capture the fact that the value of a payment instrument to a passenger depends on how widely it is accepted across routes, operators, and merchants [15-17]. Promotional Activities reflect the discount-and-cashback competition that characterizes the Indonesian mobile payment market and that operators use as a launch instrument [1], [18]. Perceived Security was measured in the source study but removed from the final model because its indicators did not separate empirically from Trust; Perceived Risk and Financial Incentives were not measured. These omissions are a boundary of the present model rather than a claim that the constructs are unimportant, and they are revisited in Section VI. Extended UTAUT models of this kind are usually estimated with Structural Equation Modeling (SEM), which reports the direction, magnitude, and statistical significance of relationships among latent variables.

SEM is the appropriate estimator when the goal is to test theory-derived hypotheses about the net effect of a given predictor on Behavioral Intention. Adoption behavior, though, may also be configurational. Intention need not rest on any single factor; it may instead depend on several favorable perceptions holding at once. Usefulness, for instance, may matter more when the system is also easy to use, socially endorsed, trusted, and widely accepted. A promotional campaign may do little on its own yet accompany higher intention once usefulness, trust, and network acceptance are already in place, a pattern consistent with recent SEM-fsQCA adoption studies that separate net effects from condition bundles [19-21]. These are directional expectations drawn from prior literature and stated here before the results are presented; they are used as interpretive orientation and not as a substitute for the formal directional expectations that an intermediate fsQCA solution would require. The additive linear specification used in the source SEM model captures such patterns only with difficulty, although SEM as a family can in principle model interactions, nonlinearity, mediation, and multigroup structures.

Configurational methods, and fuzzy-set Qualitative Comparative Analysis (fsQCA) in particular, are built for exactly this limitation. fsQCA calibrates cases into fuzzy sets and evaluates set-relational statements of sufficiency and necessity between combinations of conditions and an outcome [22]. It is not a clustering or grouping procedure: membership scores are assigned on substantive grounds, and the analysis asks whether the cases in one set are consistently contained within another. It does not require linearity, additivity, or causal symmetry. What it allows instead is equifinality, in which several recipes are associated with the same outcome; conjunctural causation, in which conditions operate in combination; and causal asymmetry, in which the configurations associated with the presence of an outcome need not be the mirror image of those associated with its absence [23-25].

Necessary Condition Analysis (NCA) adds a further and quite different logic: necessity in degree rather than in kind [26-28]. The question NCA poses is whether a condition must reach some minimum level before a given level of the outcome is observed at all, even where that condition alone can never guarantee it. The relevance to adoption research is direct, since some constructs may operate as bottlenecks rather than as drivers, a pattern documented in recent NCA work on artificial intelligence adoption, electric-vehicle sharing, and combined configurational designs [29-31]. Trust illustrates the point: it may be unable to generate adoption intention by itself, yet too little of it may coincide with an absence of high intention. It should be stressed that a ceiling line and its effect size describe an empirical boundary in cross-sectional data; they do not by themselves establish temporal or interventional causality.

The two necessity logics are not interchangeable, and the distinction matters for reading the results below. fsQCA necessity consistency asks what proportion of the outcome's set membership is contained within the condition's set membership, and it is evaluated against a benchmark such as 0.90. NCA instead estimates a ceiling in the raw XY space and quantifies the size of the empty zone above it. Because the first is a membership-containment statistic and the second is a geometric boundary statistic, the two can rank conditions differently in the same dataset. That divergence is a methodological property to be interpreted, not a contest in which one method is more valid than the other.

The distinction also matters for how findings are read substantively. Variable-centered analysis answers one question: which factors, on average, are significantly associated with adoption intention? Configurational analysis answers another: which combinations of conditions are sufficient for high adoption intention? The two are complementary rather than rival. A predictor may carry a significant SEM path and still need company before it forms a sufficient configuration; conversely, a predictor with no significant standalone estimate may still appear as a complementary component inside high-intention configurations.

The present study builds on earlier research into mobile payment intention among Trans Jatim public bus users, conducted through an extended UTAUT model [14]. That study drew on 432 cleaned survey responses and estimated the association of Performance Expectancy, Effort Expectancy, Social Influence, Trust, Network Externalities, and Promotional Activities with Behavioral Intention, with education and location tested as moderators. All predictors except Promotional Activities showed significant positive standardized estimates, ranging from 0.169 to 0.238. Those estimates are reproduced in Section III as previously reported evidence. Useful as they are for establishing the net-effect structure of adoption intention, they leave open the question of how the conditions combine, and which of them constrain the outcome.

This study therefore reanalyzes the same respondent-level dataset using fsQCA and NCA and sets the results against the previously reported SEM findings. The relationship to the prior publication is stated explicitly in Section II.A. Four research questions guide the analysis. RQ1: which conditions are significantly associated with Behavioral Intention in the previously reported net-effect baseline? RQ2: which configurations of conditions are sufficient for high Behavioral Intention, and how robust are those configurations to calibration and scoring choices? RQ3: which conditions function as necessary-in-degree constraints on Behavioral Intention, and how strong is that evidence once multiple testing is taken into account? RQ4: where do the three analytical logics converge and diverge, and what does that divergence imply for extended UTAUT?

The intended contribution is specific rather than general. It is not a new dataset, a new UTAUT specification, or a re-estimation of the original structural model. It is evidence that a predictor without a significant net effect in a well-specified SEM model, Promotional Activities, can still appear as a complementary element of sufficient configurations, and that a predictor with a mid-sized net effect, Trust, can carry the clearest necessity signal in the same data. If that pattern holds elsewhere, it implies that extended UTAUT constructs should not be evaluated on net-effect significance alone, and that non-significant extensions should not be discarded from such models without a configurational or necessity check.

II. RESEARCH METHODOLOGY

A. Research Design and Relationship to the Prior Study

This study adopts a secondary, multi-method reanalysis design. Its empirical foundation is the respondent-level survey dataset collected and analyzed by Kharismawan, Pramana, and Kristian [14] to examine mobile payment intention among Trans Jatim bus users. The present authors include the authors of [14], and the dataset, measurement items, construct definitions, and outcome variable are identical to those used there. No new respondents were recruited, no new items were administered, and no additional cases were added or removed for the present analysis.

What has already been published in [14] is the descriptive profile of the sample, the measurement assessment, the SEM path estimates, and the moderation tests. Those results are included because the fsQCA and NCA results cannot be interpreted without them. Everything else is new: the fuzzy-set calibration, the necessity and truth-table analyses, the contrarian-case diagnostic, the calibration and scoring sensitivity checks, the entire NCA component, the multiplicity assessment, and the cross-method comparison. The novelty of this article lies in applying set-theoretic and necessary-condition perspectives to previously reported net-effect evidence, not in new data collection or a new theoretical specification.

B. Sample and Data Collection

The source survey was administered online to Trans Jatim bus passengers in East Java, Indonesia. Judgmental (purposive) sampling was used to reach respondents who were active users of the service with relevant mobile payment experience; eligibility therefore required current use of Trans Jatim and familiarity with mobile payment, and the sampling frame was not a probability sample of all passengers. A total of 661 responses were collected. Of these, 229 were removed because of missing data, incomplete submission, or abnormal response patterns identified as statistical outliers, leaving 432 valid cases, a retention rate of 65.4%. Because recruitment was open and online rather than issued to a fixed list, a conventional response rate is not defined. All screening decisions were made in the original study [14]; no further exclusions were applied for the present reanalysis. The data-collection window is August 2024-March 2025.

Table 1. Respondent Profile [14]

Characteristic	Category	%
Gender	Male	61.6
	Female	38.4
Age	17-25 years	59.7
	26-35 years	25.0
	36-50 years	12.3
	>50 years	3.0
Education	Bachelor	64.8
	Non-bachelor	35.2
Residence	Surabaya	35.6
	Sidoarjo	26.2
	Gresik	16.0

Characteristic	Category	%
	Mojokerto	15.5
	Lamongan, Bangkalan, other	6.8
Mobile payment experience	1-2 years	54.9
	3-4 years	21.1
	5-6 years	16.7
	7 years or more	7.3
Instrument used*	QRIS	76.0
	GoPay	19.0
	DANA	19.0
	OVO	14.0

Table 1 summarizes the respondent profile reported in [14]. The sample is predominantly male (61.6%), young (59.7% aged 17-25), and educated to bachelor's level (64.8%), and is concentrated in the Surabaya-Sidoarjo-Gresik-Mojokerto corridor served by the network. Most respondents (76.0%) had used QRIS-based payment. These characteristics bound the external validity of both the original SEM findings and the present reanalysis: the results describe experienced, comparatively young, urban and peri-urban users of an operating service, not the general traveling public and not non-adopters. Ethical procedures for the original data collection are described in the questionnaire. It opened with an informed-consent statement describing the study, and respondents confirmed their consent before proceeding to the survey items; all participants therefore provided informed consent to participate before completing the questionnaire.

C. Measures and Previously Reported Measurement Evidence

All constructs were measured with multi-item five-point Likert scales anchored from 1 = strongly disagree to 5 = strongly agree, using items adapted from prior mobile payment and UTAUT research [7-13]. The final measurement model retained Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), Trust (TR), Network Externalities (NE), Promotional Activities (PA), and Behavioral Intention (BI). Table 2 reports the construct descriptives and Cronbach's alpha values published in [14], all of which fall between 0.850 and 0.911.

Measurement validity in the source study was assessed by exploratory factor analysis with an eigenvalue criterion of 1 and a minimum loading of 0.40, and the retained indicators met those criteria; the reported model fit was CMIN/DF = 1.701, GFI = 0.930, CFI = 0.977, and RMSEA = 0.040. Perceived Security (PS) was measured but excluded from the final model because its indicators loaded together with Trust indicators and did not display clear discriminant separation, and because its own loadings were comparatively low; discriminant validity was confirmed by factor analysis after its removal [14]. PS is retained in Table 2 only for descriptive completeness and is not used as a condition in any analysis reported here. Two limitations of this inherited evidence should be stated plainly. First, the original assessment reported Cronbach's alpha and exploratory factor analysis but did not report composite reliability, average variance extracted, or an HTMT matrix, so the exclusion of PS rests on factor-analytic rather than variance-based evidence. Second, the present study could not strengthen that evidence, because it did not re-estimate the measurement model.

Table 2. Construct Descriptives and Reliability (Previously Reported in [14])

Construct	Items	Mean	SD	Median	Cronbach's Alpha
EE	EE1, EE2, EE3	4.172	0.727	4.333	0.870
PE	PE1, PE2, PE3, PE4	4.111	0.722	4.250	0.895
SI	SI1, SI2, SI3, SI4	3.870	0.697	4.000	0.886
TR	TR1, TR2, TR3, TR4	4.036	0.730	4.000	0.893
PS	PS1, PS2, PS3, PS4	3.997	0.695	4.000	0.873
PA	PA1, PA2, PA3	4.046	0.729	4.000	0.877
NE	NE1, NE2, NE3	4.044	0.688	4.000	0.850
BI	BI1, BI2, BI3	4.021	0.842	4.000	0.911

Note. n = 432. Values reproduced from [14]. Composite reliability and average variance extracted were not reported in the original study. PS (Perceived Security) was measured but excluded from the final measurement model and is shown for descriptive completeness only; it is not used as a condition in any analysis reported here.

D. Previously Reported SEM Baseline

The SEM baseline was estimated in the original study using AMOS version 22, with PE, EE, SI, TR, NE, and PA specified as predictors of BI, and with education and location tested as moderators. This model is used here purely as a previously reported net-effect benchmark. It was not re-estimated on the present analysis file, and this article claims no new structural coefficients, fit indices, or hypothesis tests. Exact comparability between the SEM latent constructs and the composite scores calibrated below is therefore not guaranteed, and Section II.F reports a scoring sensitivity check that addresses this directly.

E. Score Construction and Fuzzy-set Calibration

For the fsQCA and NCA components, construct scores were computed at the respondent level as the unweighted arithmetic mean of the observed items belonging to each construct. All indicators were positively worded in the source

instrument and were weighted equally; no reverse coding was required, nor did the 0.500-to-0.501 membership adjustment meaningfully alter case distributions across conditions. Because cases with missing item responses had already been removed in [14], the analysis file contained no missing values, and no imputation was performed. Because the source items use a five-point Likert scale, composite means remain interpretable on the original 1-5 metric, which matters for both calibration anchors and the NCA bottleneck table.

Raw construct means were preferred to standardized z-scores because calibration requires anchors for full membership, crossover, and full non-membership that carry substantive meaning, and z-scores express distance from a sample mean rather than a level of agreement. Empirical percentile anchors were considered as an alternative and were rejected for the primary analysis for the same reason: with a sample concentrated in the upper part of the scale, percentile anchors would define 'high' relative to a favorable sample and would make the threshold move with the sample rather than with the response scale. Percentile-based anchors are nevertheless used in the factor-score sensitivity check reported in Section II.F, so that the consequences of that choice can be seen. This follows the set-theoretic principle that calibration should rest on the substantive meaning of cases rather than on statistical distribution alone [22], [24], [32].

Direct calibration used three theoretical anchors: full membership = 4.5, crossover = 3.0, and full non-membership = 1.5. A construct mean of 4.5 places the average response between 'agree' and 'strongly agree' across multiple indicators, which in this instrument is a substantively stronger marker of high favorable perception than a flat 4.0, without imposing the over-restrictive assumption that only a perfect 5.0 counts as full membership. The crossover of 3.0 is the neutral scale point, and 1.5 lies between 'strongly disagree' and 'disagree'. Cases falling at exactly 0.500 membership were adjusted to 0.501 so that truth-table allocation remained unambiguous, following standard practice [22], [25]. Because construct means are recorded to three decimals and the crossover coincides with an integer scale point, this adjustment affected only cases with a construct mean of exactly 3.000; the number of affected cases per condition ranged from 41 to 75 and is reported here for transparency.

F. fsQCA Procedure

The primary fsQCA model specified high Behavioral Intention as the outcome, with PE, EE, SI, TR, NE, and PA as causal conditions. These six conditions are fuzzy sets with continuous membership scores in the interval from 0 to 1; they are not binary variables. Six conditions generate 2^6 = 64 logically possible truth-table configurations, each corresponding to one corner of the vector space. Analysis used a frequency cutoff of three cases, a sufficiency consistency cutoff of 0.85, and a proportional reduction in inconsistency (PRI) cutoff of 0.65, all fixed before inspecting the solutions.

The frequency cutoff of three was chosen because the sample, although large by QCA standards (n = 432), is concentrated in a small number of corners: retaining rows with at least three cases keeps empirically populated but non-dominant corners in the analysis while excluding single-case corners that cannot support a consistency estimate. Because the distribution is dominated by one row, cutoffs of two or four would not materially change which corners are retained. The consistency cutoff of 0.85 is above the conventional 0.75 floor and was set a priori in view of the favorable skew of the data, on the reasoning that a lenient cutoff would be uninformative when most cases sit in high-membership regions. The PRI cutoff of 0.65 follows published guidance for avoiding simultaneous subset relations with the negated outcome [24-25], [33]. Analyses were run in fsQCA software version 4.1; logical remainders were not incorporated into the complex solution, which was derived without simplifying assumptions, while the parsimonious solution admits all remainders. No intermediate solution is reported, for the reason given below.

The analysis proceeded in six stages. Construct reliability and descriptives were inspected first, with calibration and truth-table settings summarized in Table 3. Set-theoretic necessity analysis was then conducted for each condition and its negation. Third, a contrarian-case diagnostic identified cases running against the expected positive association between favorable conditions and high BI. Fourth, truth-table sufficiency analysis was performed for both high- and low-BI outcomes. Fifth, the complex and parsimonious solutions were compared with the previously reported SEM findings. Sixth, sensitivity checks were run using stricter calibration anchors (5.0 / 4.0 / 2.0) and using factor-score proxies.

Table 3. fsQCA Calibration and Truth-table Settings

Parameter	Value	Purpose
Full membership	4.5	Strong membership in a high-condition or high-outcome set
Crossover	3.0	Point of maximum ambiguity (neutral scale point)
Full non-membership	1.5	Strong non-membership in the set
Frequency cutoff	3 cases	Minimum empirical support for a truth-table row
Consistency cutoff	0.85	Minimum sufficiency consistency, fixed a priori
PRI cutoff	0.65	Limits simultaneous subset relations with the negated outcome
Logical remainders	Not used (complex); all used (parsimonious)	Basis of the two reported solutions
Intermediate solution	Not reported	No directional expectations formally specified; core/peripheral terminology therefore not used

Note. Conditions are fuzzy sets with continuous membership scores between 0 and 1. Six conditions generate 2^6 = 64 logically possible truth-table configurations.

Two procedural decisions require explicit statement. First, this article does not report an intermediate solution and consequently does not classify any condition as core or peripheral. Core/peripheral status is established by comparing the parsimonious and intermediate solutions under stated directional expectations [25], [33], and in the absence of that comparison, the terminology would not be supported. Conditions retained under logical minimization are described here simply as 'conditions retained in the parsimonious solution'. Second, the factor-score check described next is exploratory. AMOS-estimated latent factor scores were not available in the project files, so first-component proxies were extracted from each construct's items by principal component analysis, scored by the regression method, standardized, and sign-oriented to correlate positively with the corresponding composite; these proxy scores were calibrated using empirical 95th-percentile, median, and 5th-percentile anchors, because standardized scores have no substantive 1-5 metric to which theoretical anchors could be applied. This check therefore varies score construction and calibration rule at the same time, and differences in its results cannot be attributed to either element alone. It is reported as a sensitivity analysis, not as a validation of the primary solution.

G. Contrarian-case Diagnostic

The contrarian-case diagnostic classifies cases as high or low on each condition and on BI and counts the four resulting combinations. Cases whose membership falls in an ambiguous band around the crossover are not classified in either direction. The number of unclassified cases is reported for each condition so that the denominator is visible. The membership thresholds used to define high and low were high membership = membership > 0.67 and low membership = membership < 0.33; cases with membership values from 0.33 to 0.67 were treated as intermediate membership and were not assigned to either category. The diagnostic is descriptive: it indicates whether the dataset contains cases that run against a simple monotonic reading, and it is not a test of sufficiency or necessity.

H. NCA Procedure

The NCA component used the same composite construct scores for PE, EE, SI, TR, NE, PA, and BI, and asked of each condition whether it functions as a necessary but not sufficient constraint on BI. Ceiling Envelopment with Free Disposal Hull (CE-FDH) served as the primary ceiling method. CE-FDH is the appropriate choice here because the composite scores take a limited number of discrete or semi-discrete values, so the empirical ceiling is a step function rather than a smooth line; the regression-based CR-FDH ceiling would impose linearity on that boundary and is reported alongside as a reference rather than as the primary estimate [26], [34-35]. Effect size d was computed as the empty zone above the ceiling relative to the total empirical scope, and statistical significance was assessed by a permutation test with 5,000 permutations, conducted separately for each condition. Analyses were run in SmartPLS 4.

Following common practice, d < 0.10 was read as no meaningful or negligible effect, 0.10 to < 0.30 as medium, 0.30 to < 0.50 as large, and >= 0.50 as very large; a value of exactly 0.000 is reported as no observed necessity effect rather than as a small effect. A condition was treated as empirically meaningful for necessity only where CE-FDH d >= 0.10 and p < 0.05, and where a necessity relationship was theoretically defensible. The last criterion was applied operationally and in advance: a condition was regarded as theoretically defensible for necessity if prior adoption research has argued that its absence forecloses the outcome rather than merely reducing it, which in the present model applies to Trust [10], [15], [30] and, more weakly, to Network Externalities [16]; no condition was excluded post hoc on this ground. Given the skew toward high adoption intention, this combined criterion is deliberately conservative. Bottleneck levels are reported on the original 1-5 construct scale, and the BI targets of 50% to 95% denote positions within the observed range of BI, so that a target of 90% refers to the outcome level lying 90% of the way from the lowest to the highest observed BI value, not to a percentage of respondents and not to a normalized latent score. Table 4 summarizes the specification.

Table 4. NCA Analysis Specification

Component	Specification
Outcome	Behavioral Intention (BI)
Conditions	PE, EE, SI, TR, NE, PA
Primary ceiling method	CE-FDH (step function; suited to semi-discrete scores)
Reference ceiling method	CR-FDH (reported for comparison only)
Significance test	Permutation test, 5,000 permutations, run separately per condition
Multiplicity adjustment	Holm and Benjamini-Hochberg across the six tests
Meaningful necessity criterion	CE-FDH d >= 0.10, p < 0.05, and theoretical defensibility
Bottleneck scale	Condition levels on the original 1-5 scale
Bottleneck targets	BI at 50, 60, 70, 80, 90 and 95% of the observed BI range

I. Multiple-testing Assessment

Six conditions were tested simultaneously for necessity, and the permutation p-values produced by NCA are unadjusted for that multiplicity. Because one reported result, Promotional Activities, is close to the conventional threshold, the six p-values were additionally adjusted using the Holm step-down procedure [36] and the Benjamini-Hochberg false discovery rate procedure [37]. The adjusted values are used throughout the process to grade the strength of the necessity evidence.

Six conditions were tested simultaneously for necessity, and the permutation p-values produced by NCA are

unadjusted for that multiplicity. Because one reported result, Promotional Activities, is close to the conventional threshold, the six p-values were additionally adjusted using the Holm step-down procedure [36] and the Benjamini-Hochberg false discovery rate procedure [37]. The adjusted values are used throughout the process to grade the strength of the necessity evidence.

III. RESULTS

A. Previously Reported Net-Effect Baseline

The previously reported SEM findings supply the variable-centered baseline (Table 5). Five predictors, namely Performance Expectancy, Effort Expectancy, Social Influence, Trust, and Network Externalities, showed significant positive standardized estimates for Behavioral Intention, ranging from 0.169 (EE) to 0.238 (PE), while Promotional Activities showed no significant standalone estimate (0.129, p = 0.077). These are modest standardized effects: with n = 432, statistical significance should not be read as evidence of large substantive importance, and confidence intervals were not reported in the original study. Perceived utility, ease, social endorsement, trust, and network acceptance are therefore associated on average with intention, whereas promotions on their own are not an independent net-effect pathway in this specification. Hypothesis numbering follows the original study, in which H5 concerned Perceived Security; that construct was removed from the final measurement model, and H5 is therefore neither listed nor tested here, as noted beneath Table 5.

Table 5. Previously Reported SEM Baseline [14]

Hypothesis	Path	Standardized Estimate	p-value	Conclusion
H1	PE -> BI	0.238	0.004	Supported
H2	EE -> BI	0.169	0.017	Supported
H3	SI -> BI	0.210	0.005	Supported
H4	TR -> BI	0.219	0.009	Supported
H6	NE -> BI	0.179	0.026	Supported
H7	PA -> BI	0.129	0.077	Not supported

Note. n = 432. Standardized regression weights estimated in AMOS 22 and reported in [14]; not re-estimated in the present study. Confidence intervals were not reported in the original study. Hypothesis numbering follows [14]: H5 concerned Perceived Security, which was removed from the final measurement model and is therefore not tested. Model fit reported in [14]: CMIN/DF = 1.701, GFI = 0.930, CFI = 0.977, RMSEA = 0.040.

B. Set-theoretic Necessity

Necessity analysis in fsQCA returned high necessity consistency for most conditions (Table 6). Measured against the conventional 0.90 benchmark, PE, EE, TR, NE, and PA all exceeded it, and SI met it at 0.900. These values require a cautious reading, and the 0.90 figure should be treated as a heuristic benchmark rather than a substantive threshold. The dataset is strongly skewed toward favorable responses and high Behavioral Intention: under the primary calibration, 404 of 432 cases (93.5%) have greater membership in than out of the high-BI set. When most cases sit in high-membership regions on both the condition and the outcome, containment of the outcome set within the condition set becomes almost arithmetically likely, and necessity consistency is inflated accordingly. The necessity coverage values in Table 6 are correspondingly high across all conditions and do not discriminate between them. A relevance-of-necessity diagnostic would sharpen this assessment and could not be computed from the available output; it is noted as a limitation in Section VI. For these reasons, no condition is claimed here as substantively necessary on the basis of Table 6 alone, and the NCA results reported in Section III.F are used as the primary evidence on necessity.

Table 6. fsQCA Necessity Analysis for High BI

Condition	Necessity Consistency	Necessity Coverage
EE	0.948	0.905
PE	0.947	0.918
TR	0.936	0.926
NE	0.932	0.916
PA	0.926	0.913
SI	0.900	0.933
~SI	0.236	0.885
~NE	0.195	0.914
~PA	0.194	0.896
~TR	0.189	0.860
~PE	0.169	0.850
~EE	0.159	0.866

Note. The 0.90 figure is a heuristic benchmark, not a substantive threshold. Because 404 of 432 cases (93.5%) hold greater membership in than out of the high-BI set, necessity consistency is inflated for all conditions and does not discriminate among them. No condition is claimed as substantively necessary on the basis of this table.

C. Sufficiency: Truth Table and Solutions

Under the primary thresholds, the truth-table analysis returned seven rows consistent with high BI. Empirical

diversity is severely limited: the all-present configuration alone contains 380 of the 432 cases, that is 88.0% of the sample, and the remaining populated rows contain small numbers of cases. This concentration is itself a substantive diagnostic finding and not merely a caveat. When one corner of the vector space holds almost nine cases in ten, logical minimization has little cross-corner variation to work with, and solutions in which any five of the six conditions appear can emerge close to mechanically from the dominance of the all-present row. The complete truth table, including configuration, case count, raw consistency, PRI, and outcome coding for all populated rows, is available from the authors on request and should ideally accompany the article as supplementary material.

Logical minimization reduced the seven rows to six solution paths, each combining five of the six conditions, with solution consistency 0.946 and solution coverage 0.904 (Table 7). The accurate statement of this result is that six observed five-condition configurations met the sufficiency criteria for membership in the high-BI set at the specified calibration and truth-table thresholds. It is not that high BI is guaranteed whenever any five conditions are present. The six paths overlap extensively: raw coverage is between 0.798 and 0.820 for every path, while unique coverage ranges only from 0.007 to 0.029. In other words, each path covers substantially the same cases as the others, and almost no case is explained by one path alone. Substantive equifinality is therefore weak. What the solution mainly records is that removing any single condition from the all-present bundle still leaves a configuration consistent with high BI in these data.

Table 7. Complex Solution for High BI (Primary Calibration)

Solution Path	Consistency	Raw Coverage	Unique Coverage
EE*SI*TR*NE*PA	0.977	0.807	0.017
PE*SI*TR*NE*PA	0.977	0.812	0.021
PE*EE*TR*NE*PA	0.966	0.820	0.029
PE*EE*SI*NE*PA	0.972	0.798	0.007
PE*EE*SI*TR*PA	0.978	0.810	0.019
PE*EE*SI*TR*NE	0.978	0.811	0.021
Overall solution	0.946	0.904 (solution coverage)	-

Note. '*' denotes logical AND. Seven truth-table rows met the sufficiency criteria; the all-present row alone contains 380 of 432 cases (88.0%). Raw coverage is almost identical across paths while unique coverage never exceeds 0.029, so the six paths describe largely the same cases and substantive equifinality is limited.

Table 8. Parsimonious Solution for High BI (Primary Calibration)

Solution Term	Consistency	Raw Coverage	Unique Coverage
EE	0.905	0.948	0.035
PE	0.918	0.947	0.034
Overall solution: EE + PE	0.894	0.982 (solution coverage)	-

Note. '+' denotes logical OR. EE and PE are two alternative single-condition parsimonious terms, not a joint PE*EE configuration (Table 8); the first two rows report each term separately, and the third reports the union of the two. The parsimonious solution incorporates all logical remainders and is deliberately permissive. Because no intermediate solution is reported, neither term is described as a core condition.

The parsimonious solution is EE + PE, with solution consistency 0.894 and solution coverage 0.982 (Table 8). In Boolean notation, the '+' operator denotes logical OR, not conjunction, which is written '*'. The parsimonious solution therefore consists of two alternative single-condition terms: high EE is sufficient for high BI in these data, and, separately, high PE is sufficient for high BI. It is not a joint PE*EE configuration, and the two conditions are not required to occur together. Each term performs comparably on its own, with consistency 0.905 and raw coverage 0.948 for EE and consistency 0.918 and raw coverage 0.947 for PE, and each contributes only about 0.035 in unique coverage. This does not render Social Influence, Trust, Network Externalities, or Promotional Activities irrelevant; parsimonious solutions admit all logical remainders and are deliberately permissive. It means that, under maximally permissive minimization, either usefulness or ease of use alone suffices to reproduce the high-BI cases in this dataset. Because no intermediate solution is reported, neither condition is described as core.

No low-BI configuration cleared the same sufficiency thresholds. That absence should not be read as evidence that low adoption intention has no causes. Under the primary calibration, only 28 of 432 cases (6.5%) have greater membership out of than in the high-BI set, and under the stricter calibration reported below, the figure rises only to 106 cases. With so few cases in the negated outcome, populated truth-table rows for ~BI are too sparse to support consistency estimates at a 0.85 cutoff, and applying the same threshold to such an unbalanced outcome would in any case be questionable. The analysis is therefore informative about the

configurations associated with high intention and uninformative about the configurations associated with its absence.

The contrarian-case diagnostic is consistent with this reading (Table 9). Every condition produced a small number of high-condition/low-BI and low-condition/high-BI cases, so the dataset is not perfectly deterministic, but these cases are sparse next to the high-condition/high-BI cell, which holds between 245 and 294 cases depending on the condition. Between 119 and 172 cases per condition fall in the ambiguous middle band and are not classified, and those counts are now reported in the table. The empirical setting is best characterized as a favorable, comparatively mature adoption-intention context rather than a balanced sample of adopters and non-adopters.

Table 9. Contrarian-case Diagnostic

Condition	High cond. / low BI	Low cond. / high BI	High cond. / high BI	Low cond. / low BI	Classified	Unclassified
PE	9	4	294	6	313	119
EE	15	5	286	6	312	120
SI	7	1	245	7	260	172
TR	9	1	282	8	300	132
NE	16	5	258	4	283	149
PA	11	4	252	4	271	161

Note. n = 432 for every condition. Cases whose membership falls in the ambiguous band around the crossover are not classified in either direction and are reported in the final column, which is why the four classification cells do not sum to 432. Membership thresholds used to define high and low are stated in Section II.G.

D. Calibration Sensitivity

Calibration sensitivity was assessed by re-running the analysis with stricter anchors (Table 10). Moving full membership to 5.0, crossover to 4.0, and full non-membership to 2.0 does more than tighten the calibration: it relocates the point of maximum ambiguity from the neutral scale midpoint to 'agree', so that a respondent who agrees with every item is no longer more in than out of the set. That is a materially different substantive definition of a high condition, and it was chosen precisely because it tests whether the primary findings depend on a permissive definition of 'high' in a sample whose modal response is 4. Under that definition, the high-BI set falls from 404 to 326 cases and the all-present row from 380 to 149, while complex solution consistency falls from 0.946 to 0.892 and coverage from 0.904 to 0.724.

These figures show that model fit deteriorates but remains above conventional consistency thresholds. They do not, on their own, establish that the same recipes survive, because Table 10 reports fit statistics and case counts rather than the minimized solution expressions. The appropriate conclusion is therefore narrower than a claim of structural robustness: the sufficiency relationship between favorable conditions and high BI persists under a stricter definition of set membership, with visibly weaker empirical support, while the stability of the specific configurations under that calibration is not demonstrated by the evidence presented here.

Table 10. Calibration Sensitivity

Scenario	Anchors (In / Cross / Out)	High BI Cases	All-Present Row Cases	Complex Consistency	Complex Coverage	Solution Expression
Primary (theoretical Likert anchors)	4.5 / 3.0 / 1.5	404	380	0.946	0.904	Six five-condition paths
Strict (high-agreement anchors)	5.0 / 4.0 / 2.0	326	149	0.892	0.724	Not reported

Note. The strict scenario moves the crossover from the neutral scale point to 'agree', which redefines set membership substantively rather than merely tightening it. Because the minimized solution expression under the strict scenario is not reported, this table shows that model fit deteriorates but does not establish that the same configurations survive.

E. Scoring Sensitivity

The factor-score proxy check produces a materially different configurational picture, and the difference is reported here rather than minimized. Its complex solution retains acceptable sufficiency consistency at 0.906 with coverage 0.725, but the solution terms are not those of the primary analysis, and three of the seven terms contain negated conditions: EE*~SI*NE*PA, PE*~SI*TR*PA, PE*EE*~SI*PA, and PE*EE*~SI*TR*NE all include the absence of Social Influence, while PE*EE*SI*TR*~NE includes the absence of Network Externalities. The proxy parsimonious solution is PE*TR + SI*TR*PA + EE*NE*PA + PE*EE*PA, which is a set of four conjunctive terms rather than the two alternative single-condition terms of the primary analysis.

Two conclusions follow. The first is general and holds across specifications: high BI is associated with combinations of conditions rather than with any single isolated condition, and Trust and Promotional Activities recur across the proxy terms. The second is specific and does not hold: the particular recipes are not stable across scoring and calibration choices, and the prominence of PE and EE in the primary parsimonious solution is not reproduced when scores are constructed and calibrated differently. Because this check alters score construction and calibration rules simultaneously, it cannot attribute the instability to either factor, and it is reported as an exploratory sensitivity analysis (Table 11). The appearance of negated conditions also indicates configurational asymmetry that the primary specification, with its all-present dominant row, has little power to detect. Claims about PE and EE as configurational anchors are moderated accordingly in Sections IV and V.

Table 11. Factor-score Proxy Sensitivity Analysis (Exploratory)

Solution Type	Positive Rows	Solution	Consistency	Coverage
Complex	16	SI*TR*NE*PA + EE*~SI*NE*PA + PE*~SI*TR*PA + PE*EE*NE*PA + PE*EE*~SI*PA + PE*EE*~SI*TR*NE + PE*EE*SI*TR*~NE	0.906	0.725
Parsimonious	16	PE*TR + SI*TR*PA + EE*NE*PA + PE*EE*PA	0.866	0.791

Note. '~' denotes the absence of a condition. This check alters score construction (principal component proxy scores) and calibration rule (95th percentile / median / 5th percentile anchors) simultaneously, so differences cannot be attributed to either element alone. The parsimonious solution differs materially from the primary result (EE + PE), and four terms contain negated conditions. It is reported as an exploratory sensitivity analysis, not as validation.

Table 12. NCA Effect Sizes and Permutation Tests

Condition	Outcome	CE-FDH d	CR-FDH d	p (unadjusted)	Effect Size	Meaningful necessity
TR	BI	0.208	0.186	< 0.001	Medium	Yes
SI	BI	0.125	0.087	0.184	Medium	No (not significant)
PA	BI	0.104	0.091	0.044	Medium	Nominal only (see Table 13)
PE	BI	0.062	0.069	0.563	Below threshold	No
EE	BI	0.042	0.029	0.501	Below threshold	No
NE	BI	0.000	0.000	1.000	No observed effect	No

Note. n = 432. Effect-size bands: d < 0.10 negligible or below the meaningfulness threshold, 0.10 to < 0.30 medium, 0.30 to < 0.50 large, >= 0.50 very large; d = 0.000 denotes no observed necessity effect rather than a small one. p-values in this table are unadjusted for the six simultaneous tests; adjusted values are given in Table 13.

F. Necessary Condition Analysis

NCA identifies a different form of empirical relevance (Table 12). Trust returned the largest observed CE-FDH necessity effect among the tested conditions, d = 0.208 with p < 0.001, a medium effect by the conventional benchmark. Social Influence returned a medium effect size of d = 0.125 but a permutation p-value of 0.184, so its effect size cannot be distinguished from what unrelated variables of similar distribution would produce; effect-size magnitude and evidence against randomness are separate matters, and SI is not interpreted as a bottleneck anywhere in this article. Promotional Activities returned d = 0.104 with p = 0.044, marginally above the effect-size criterion and marginally below the conventional significance level. Performance Expectancy and Effort Expectancy returned small effects that were not significant, and Network Externalities returned d = 0.000, that is, no observed necessity effect at all rather than a small one.

Because six conditions were tested simultaneously, the unadjusted p-values overstate the evidence. Table 13 reports Holm-adjusted and Benjamini-Hochberg-adjusted values. Trust remains significant under both procedures. Promotional Activities does not: its Holm-adjusted p-value is 0.220, and its Benjamini-Hochberg-adjusted p-value is 0.132, both well above 0.05. The necessity evidence in this study is therefore of two clearly unequal grades. Trust carries robust evidence of a necessary-in-degree relationship with Behavioral Intention within this dataset. Promotional Activities carries nominal, borderline evidence that does not survive correction for multiplicity, and every statement about PA as a necessary or minimum enabling condition in this article is qualified accordingly.

This ordering diverges from the fsQCA necessity consistencies, where PE and EE ranked highest. The divergence is a methodological property rather than a contradiction, and neither ranking is simply more correct. fsQCA necessity consistency measures containment of outcome membership within condition membership and is inflated when most cases occupy high-membership regions of both sets, which is the case here. NCA estimates a ceiling in the raw XY space and quantifies the empty zone above it, so it responds to the shape of the upper-left boundary rather

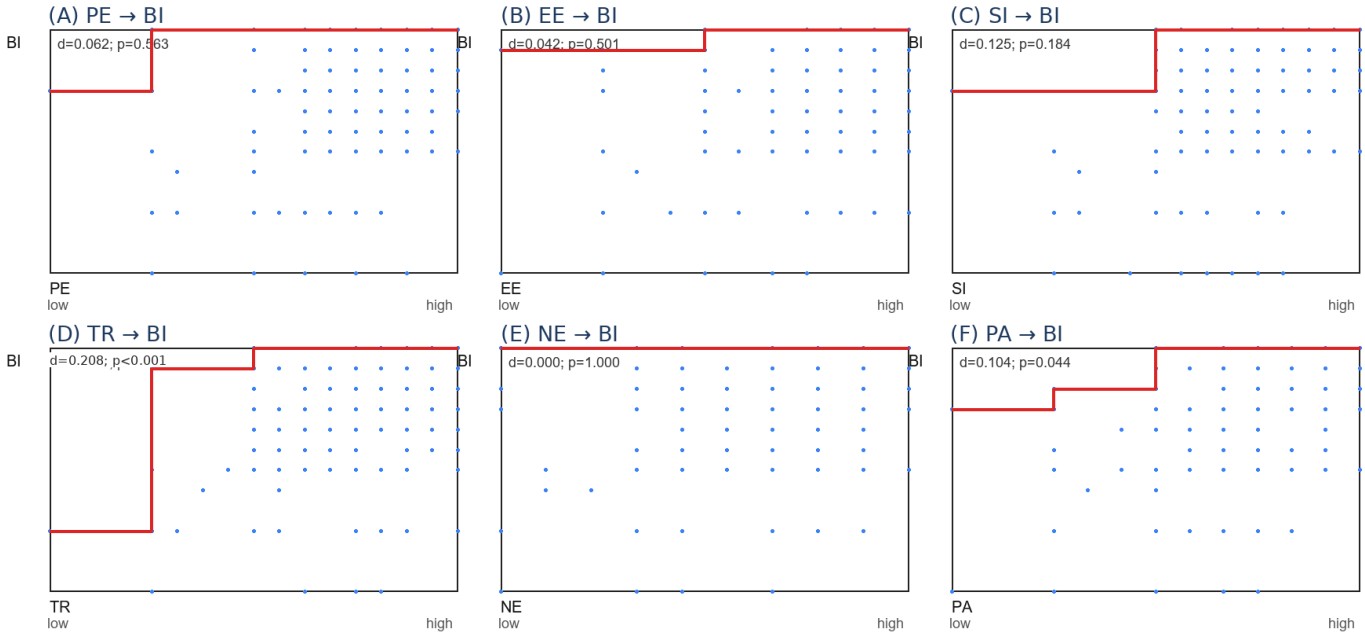
than to the mass of cases. That makes NCA the more informative diagnostic for necessity in a favorably skewed sample such as this one. It does not make NCA immune to distributional problems: range restriction, floor and ceiling effects, and a narrow spread of observed condition values all affect the size of the empty zone and hence d, and the present sample is restricted in exactly that way.

Table 13. Multiplicity-adjusted Permutation p-values for the Six NCA Tests

Condition	CE-FDH d	p (unadjusted)	p (Holm)	p (Benjamini-Hochberg)	Significant at 0.05 after adjustment
TR	0.208	< 0.001	< 0.006	< 0.006	Yes (both procedures)
PA	0.104	0.044	0.220	0.132	No
SI	0.125	0.184	0.736	0.368	No
EE	0.042	0.501	1.000	0.676	No
PE	0.062	0.563	1.000	0.676	No
NE	0.000	1.000	1.000	1.000	No

Note. Adjustments computed across the six condition-level tests reported in Table 12, using the Holm step-down procedure [36] and the Benjamini-Hochberg false discovery rate procedure [37]. Only Trust remains significant after adjustment.

The ceiling plots in Figure 1 illustrate the same results visually and are a diagnostic of that analysis rather than independent confirmation of it, since they are generated from the same data by the same method. In each panel, the horizontal axis carries the condition and the vertical axis Behavioral Intention, both increasing to the right and upward; the red step function is the CE-FDH ceiling drawn along the upper-left edge of the observed data cloud, and d is computed from the area of the empty zone above that line relative to the total empirical scope. The visibly largest empty zone occurs for Trust (panel D), with smaller zones for Promotional Activities (panel F) and Social Influence (panel C), and essentially none for Network Externalities (panel E).



Note. (A) Performance Expectancy, (B) Effort Expectancy, (C) Social Influence, (D) Trust, (E) Network Externalities, (F) Promotional Activities. In each panel, the horizontal axis carries the condition and the vertical axis carries Behavioral Intention, both increasing rightward and upward. The red step function is the CE-FDH ceiling; the empty zone above it is the basis of the effect size d. Reported p-values are unadjusted permutation p-values; adjusted values are given in Table 13.

Figure 1. NCA Ceiling Plots (CE-FDH) for the Six Conditions Against Behavioral Intention.

The bottleneck table (Table 14) reports the minimum observed condition levels associated with selected BI targets on the original 1-5 construct scale, where each target denotes a position within the observed range of BI. Required levels stay low across the lower targets and rise only in the upper range. At the 90% target, the minimums are TR = 2.0, PA =

3.0, SI = 3.0, PE = 2.0, and NE = 2.0; at 95%, both TR and EE rise to 3.0 while SI and PA remain at 3.0. Three qualifications apply. First, the table lists all conditions for completeness, but only Trust satisfies the necessity criteria used in this study once multiplicity is taken into account, and the rows for SI, PE, EE, NE and PA should not be read as established necessary conditions; Social Influence in particular imposes a numerically high floor at the upper targets despite a non-significant permutation test. Second, these values are empirical lower bounds observed within this sample's range, not managerial targets: NCA identifies what constrains the outcome, not what secures it. Third, because they are bounded by the observed data, they would change with a sample containing more low-intention respondents.

Table 14. NCA Bottleneck Table (Condition Levels on the Original 1-5 Scale)

Condition	BI 50%	BI 60%	BI 70%	BI 80%	BI 90%	BI 95%
TR †	2.00	2.00	2.00	2.00	2.00	3.00
PA ‡	1.00	1.00	1.00	2.00	3.00	3.00
SI ‡	1.00	1.00	1.00	3.00	3.00	3.00
PE ‡	1.00	1.00	1.00	2.00	2.00	2.00
EE ‡	1.00	1.00	1.00	1.00	1.00	3.00
NE ‡	2.00	2.00	2.00	2.00	2.00	2.00

Note. BI targets denote positions within the observed range of Behavioral Intention: BI 90% is the outcome level 90% of the way from the lowest to the highest observed BI value. Values are empirical lower bounds observed within this sample, not managerial targets, and they depend on the sample's observed range. † Condition meeting the necessity criteria after multiplicity adjustment. ‡ Condition not established as necessary in this study; its bottleneck values should not be read as required levels.

IV. DISCUSSION

The previously reported SEM baseline and the present fsQCA reanalysis address mobile payment intention in complementary ways. SEM identifies predictors with statistically significant average associations with Behavioral Intention; fsQCA identifies combinations of favorable perceptions that are sufficient for membership in the high-intention set. The two answer different questions rather than competing to answer the same one, a division of labor also drawn in recent hybrid adoption studies that pair SEM for net effects with fsQCA for alternative sufficient configurations [19-20]. Throughout this section, 'causal' is used in the sense of the causal logics these methods operationalize, namely set-relational sufficiency and necessity; the cross-sectional observational design does not establish temporal or interventional causality.

Convergence between the two logics is real but narrower than it first appears. Performance Expectancy and Effort Expectancy were significant predictors in the SEM baseline, and they are also the conditions retained in the parsimonious fsQCA solution under the primary specification. That solution, however, is EE + PE, two alternative minimized terms, so the correct reading is that either condition alone reproduces the high-BI cases under maximally permissive minimization, not that the two jointly constitute a configuration. The convergence is further qualified by the scoring sensitivity check, which returned an entirely different parsimonious solution. PE and EE are therefore best described as prominent in the primary specification rather than as reliable or central conditions. For a public transport payment system, the defensible statement is that passengers in this sample reported high intention where mobile payment was seen as worthwhile or as easy to use, a pattern in line with mobile payment studies placing usefulness, ease, and trust among the central antecedents of adoption [9-10], [12-13].

fsQCA also extends the net-effect reading by foregrounding conjunctural patterns, though the strength of that extension is bounded by the data. In the complex solution, high BI is associated with configurations in which nearly every favorable condition is present. Social influence, trust, network externalities, and promotional activities are better understood as conditions that co-occur within sufficient configurations than as free-standing variables, a conclusion echoed in configurational studies showing adoption outcomes arising from several distinct condition packages [19-20], [38]. Co-occurrence within configurations is not, however, evidence of an interaction mechanism, which was not tested. Nor is the observed equifinality substantial: with unique coverage between 0.007 and 0.029 and one truth-table row holding 88% of the cases, the six paths are near-redundant descriptions of the same dominant configuration rather than six distinct routes to high intention.

NCA adds a third layer by separating conditions that are associated with the outcome on average from conditions that appear to bound it. Trust was a significant predictor in the SEM baseline, appears in the high-BI configurations, and returns the largest observed necessity effect. Within this dataset, Trust therefore shows the clearest evidence of a bottleneck relationship with behavioral intention, with a medium effect size of d = 0.208 and significance surviving adjustment for multiple testing. High intention does not follow from trust alone, but low trust coincides with an absence of high intention across this sample, a reading compatible with mobile payment research on security, risk, and trust and with NCA evidence that trust can act as a critical threshold in mobility service adoption [10], [15], [30]. This is an empirical property of the present sample and specification, not a universal property of the construct. The distinction from SEM remains useful: a standardized coefficient describes average association, whereas a ceiling locates the level below which the outcome is not observed in these data.

The sharpest divergence concerns Promotional Activities, and it is here that the three grades of evidence must be kept apart. PA showed no significant standalone estimate in the previously reported SEM model (0.129, p = 0.077), which from a variable-centered standpoint means promotions did not predict intention once other constructs were accounted for. PA nevertheless appears as a component of several sufficiently high-BI configurations, and it returns a nominally significant NCA necessity effect that does not survive multiplicity adjustment. These are three findings of quite unequal strength: a clean non-result in a well-fitting structural model, a robust but weakly discriminating configurational presence, and a borderline necessity signal. Taken together, they suggest that PA did not show an independent net-effect association but appeared as a complementary condition within configurations, which marks a boundary condition for extending UTAUT in public transport payment contexts. Similar patterns appear in adoption studies where policy support, promotion, or contextual incentives operate through broader adoption readiness rather than as universal direct effects [1], [18-19]. The theoretical claim that survives this evidence is modest: in this dataset and context, promotional incentives behaved as configurational complements rather than as independent antecedents of intention.

The NCA result for PA should be read in the same key. It is tentative rather than confirmatory: d = 0.104 is barely above the meaningfulness threshold, the unadjusted p-value of 0.044 is marginal, and neither the Holm nor the Benjamini-Hochberg adjustment leaves it significant. Read with those qualifications, the evidence is compatible with promotional activity functioning as a minimum enabling condition within a wider adoption package, within the observed empirical scope, but it does not establish that it does. What can be said without qualification is the negative finding: promotional activity was not an independent driver of intention in this sample.

This pattern illustrates what is gained by supplementing symmetric net-effect analysis with asymmetric and conjunctural logics. The SEM baseline used here specified additive linear net effects and assumed that the same model structure describes high and low intention alike. fsQCA relaxes both assumptions, permitting several recipes for the same outcome and accepting that the absence of high BI need not be explained by reversing what accompanies it [24], [31]. In this dataset, the missing low-BI configurations mark an empirical boundary of the sample rather than a theoretical finding: the sample is concentrated at high intention levels, so the study is equipped to describe how favorable intention is configured and not why resistance forms.

That six five-condition combinations meet the sufficiency criteria means high intention was not observed to require every favorable condition simultaneously. Given the extensive overlap among those combinations, however, the more defensible reading is that high intention in this sample accompanies a dense bundle of favorable conditions from which any one element can be missing. NCA then qualifies the picture by indicating that one condition, Trust, bounds the outcome more tightly than the others. Interpreted for this dataset, the extended UTAUT model behaves neither as a set of independent additive paths nor as a set of sharply distinct sufficient recipes, but as a bundle within which conditions play different roles: some are retained under minimization, some appear only as complements, and one operates as a constraint. Whether that role assignment is a stable property of these constructs, or an artifact of this sample's favorable distribution, cannot be settled here and requires replication in samples with greater variation in intention.

V. MANAGERIAL IMPLICATIONS

For Trans Jatim and comparable operators, the combined evidence points toward a bundled approach to building adoption intention. The recommendations below are implications drawn from cross-sectional associations, set relations, and ceiling constraints; none of them has been tested as an intervention in this study. Performance Expectancy and Effort Expectancy are the conditions retained in the primary parsimonious solution, and each was independently associated with high intention, so investment in perceived usefulness and perceived ease of use is the most directly supported priority. In practice, this concerns interface simplicity, transaction reliability, clear instructions, and low-friction QR or wallet use at the point of boarding. This is a recommendation about perceptions of the mobile payment channel itself; the study did not compare mobile payment with cash experimentally and makes no claim that one is faster or more convenient than the other. The point carries additional weight because cashless transit can raise practical barriers where implementation overlooks passenger capability, access, and payment reliability [4-5], [39].

Trust may be treated as a bottleneck, with the qualification that the underlying effect is medium in size (d = 0.208) and that the thresholds observed are specific to this sample. The NCA results indicate that high intention was not observed where trust was low, however favorable the other conditions. Operators can address trust through transaction reliability, clear refund and failure-handling procedures, visible institutional backing, and honest communication when transactions fail or stall. One concrete way to act on this, offered as a suggestion rather than as a demonstrated policy, is to treat a baseline level of transaction reliability as a precondition for scaling promotional campaigns: an operator might, for example, set an internal target for successful-transaction rate and passenger-reported payment confidence, and defer discount or cashback campaigns until that target is met, on the reasoning that promotional spending is unlikely to compensate for a payment channel that passengers do not yet trust. The present data support the ordering implied by that suggestion but do not establish its effectiveness.

Social Influence and Network Externalities co-occur with the other conditions in the sufficient configurations, which suggests they are best considered alongside the rest of the adoption package rather than in isolation. This is an interpretive suggestion rather than an organizational prescription; co-occurrence within configurations does not by itself demonstrate that communication programs should be

merged. Social influence can be activated by peer normalization, guidance from drivers or station staff, and passenger-facing cues that make mobile payment use observable. Network externalities strengthen as payment compatibility broadens and acceptance becomes predictable across routes and service contexts, and recent research suggests that network effects and trust facilitators shape both initial acceptance and continuance intention [15-16]. It should be noted that Social Influence returned a non-significant NCA permutation test and is not treated here as a bottleneck.

Promotional Activities call for caution rather than dismissal, and the caution now runs in both directions. The SEM baseline found no independent association between promotions and intention, while fsQCA places PA inside sufficient configurations and NCA returns only borderline, non-robust necessity evidence for it. The defensible managerial reading is therefore that discounts, cashback, and campaign incentives should not be assumed to compensate for deficiencies in usefulness, ease of use, or trust. They are more defensible as short-term reinforcement once those conditions are in place, which matches cashless-payment studies treating policy support, campaigns, and adoption readiness as parts of one enabling environment [1], [18-19]. The evidence here does not support presenting PA as a condition with the same evidential standing as Trust.

VI. LIMITATIONS AND FUTURE RESEARCH

Several limitations qualify these conclusions. The first and most consequential is conceptual: the outcome is self-reported Behavioral Intention, not observed mobile payment transactions or actual use. Intention is an established antecedent of behavior but is not equivalent to it, and none of the findings reported here should be read as evidence about realized adoption, transaction volumes, or sustained use. Self-reported measures of this kind, collected in a single instrument, are also exposed to common-method variance, which the source study did not test for and which could inflate the observed associations among constructs.

The second limitation concerns empirical diversity. The sample is strongly concentrated at high intention and high agreement, with 380 of 432 cases in the all-present truth-table row and only 28 cases holding greater membership out of than in the high-BI set under the primary calibration. This concentration has direct analytical consequences rather than being a general caveat. It limits the cross-corner variation available to logical minimization, which is why the sufficient configurations overlap so heavily and why their unique coverage is negligible; it inflates fsQCA necessity consistency across all conditions; it makes the absence of any low-BI solution uninformative about the conditions accompanying low intention; and it restricts the range over which NCA ceilings are estimated, so the reported bottleneck levels reflect the observed range of this sample. The study should therefore be read as an account of how high intention is configured in a favorable public transport context, not as a general account of adoption or of resistance.

A third limitation concerns the design as a whole. This is a secondary reanalysis of an existing cleaned dataset and inherits the measurement decisions, screening rules, and sampling frame of the prior research [14], including judgmental sampling of active users, the absence of composite reliability and average variance extracted statistics, and the exclusion of Perceived Security on factor-analytic grounds. Perceived Risk and Financial Incentives, which the literature identifies as relevant to mobile payment, were not measured and could not be entered as conditions. Because the design is cross-sectional, no temporal ordering among the constructs can be established, and the configurational and ceiling relationships reported here describe co-occurrence and empirical boundaries rather than intervention effects.

Another limitation concerns measurement and scoring. The primary fsQCA and NCA analyses use construct means rather than latent factor scores exported from the validated SEM model. Composite scores were chosen because they preserve the meaning of the 1-5 response scale for calibration and bottleneck interpretation, and the factor-score proxy check was added to reduce reliance on arithmetic means. That check changed score construction and calibration rule simultaneously and returned a different parsimonious solution, which shows that the specific recipes are sensitive to these choices. The authors regard a clean separation of scoring sensitivity from calibration sensitivity, using AMOS-exported latent scores under identical anchors, as the single most valuable next step; it could not be carried out here because the latent scores were not available in the project files. Until that comparison is made, claims about which particular conditions are retained under minimization should be treated as provisional.

Future research should address these limitations directly. Validating the present findings against observed transaction data, and against longitudinal measurement that separates intention from subsequent use, would test the conceptual boundary that constrains this study most tightly. Collecting additional data from infrequent users, cash-preferring passengers, and non-adopters would supply the empirical diversity that the present sample lacks and would allow the asymmetry of configurational analysis to be exploited properly. Subgroup fsQCA and NCA by travel origin, education, age, or prior mobile payment experience offer a further avenue, as does triangulating SEM, fsQCA, NCA, and predictive modeling to separate explanatory, configurational, bottleneck, and forecasting value in adoption research [17], [21], [24], [31], [40].

VII. CONCLUSION

This study reanalyzed the respondent-level dataset of a previously published extended UTAUT study of mobile payment intention among Trans Jatim bus passengers, applying fsQCA and NCA and reading the results against the SEM estimates reported in that earlier work. The SEM baseline, which was not re-estimated here, showed significant positive associations between Performance Expectancy, Effort Expectancy, Social Influence, Trust, and Network

Externalities and Behavioral Intention, with Promotional Activities showing no significant standalone association. fsQCA extends that picture: at the stated thresholds, high Behavioral Intention is associated with configurations of favorable perceptions, the complex solution resolving into six heavily overlapping five-condition paths, and the parsimonious solution consisting of two alternative single-condition terms, EE or PE. NCA identifies Trust as the condition with the largest observed necessity effect, robust to multiplicity adjustment, and Promotional Activities as a nominal and non-robust one.

The contributions are methodological and theoretical, and both are bounded by the evidence. Methodologically, the study shows what combining previously reported net-effect evidence with configurational and necessary-condition analysis can add, and equally what it cannot: with a sample concentrated in one corner of the vector space, sufficiency solutions are near-redundant, and the absence of a negated-outcome solution is uninformative. Theoretically, the finding that matters is the asymmetry of evidence across methods for a single construct. In this dataset and context, Promotional Activities was not supported as an independent predictor of intention, yet appeared as a complementary condition within sufficient configurations; the necessity evidence for it is tentative and does not survive correction for multiple testing. The implication for extended UTAUT is that a non-significant extension variable should not be discarded on net-effect grounds alone without a configurational or necessity check, and that Trust, in settings of this kind, may warrant treatment as a constraint on intention rather than only as one predictor among several. Both propositions are offered as interpretations of a single cross-sectional reanalysis and require replication in samples with greater variation in the outcome.

DECLARATION

During the preparation of this work, the author(s) used Grammarly for grammar checking and Google Gemini to improve language clarity, readability, and the presentation of revisions made in response to reviewer comments. No part of the study design, data analysis, results, or interpretation was generated by these tools. The author(s) reviewed and edited all content and take full responsibility for the content of the published article.

REFERENCES

[1] Balakrishnan V, Shuib NLM. Drivers and inhibitors for digital payment adoption using the Cashless Society Readiness-Adoption model in Malaysia. *Technology in Society*. 2021;65:101554. doi:10.1016/j.techsoc.2021.101554.

[2] Rahman M, Ismail I, Bahri S. Analysing consumer adoption of cashless payment in Malaysia. *Digital Business*. 2021;1(1):100004. doi:10.1016/j.digbus.2021.100004.

[3] Yang M, Al Mamun A, Mohiuddin M, Che Nawi N, Zainol NR. Cashless Transactions: A Study on Intention and Adoption of e-Wallets. *Sustainability*. 2021;13(2):831. doi:10.3390/su13020831.

[4] Golub A, Brown A, Brakewood C, MacArthur J, Lee S, Ziedan A. Equity and exclusion issues in cashless fare payment systems for public transportation. *Transportation Research Interdisciplinary Perspectives*. 2022;15:100628. doi:10.1016/j.trip.2022.100628.

[5] Azmoodeh M, Sheykhfard A, Tabatabaei F, Azimi M, Guo QV, Soltani N. Modeling Barriers to Cashless Fare Adoption in U.S. Public Transit Systems. *Transportation Research Record: Journal of the Transportation Research Board*. 2026. doi:10.1177/03611981261431737.

[6] Khanra S, Dhir A, Kaur P, Joseph RP. Factors influencing the adoption postponement of mobile payment services in the hospitality sector during a pandemic. *Journal of Hospitality and Tourism Management*. 2021;46:26-39. doi:10.1016/j.jhtm.2020.11.004.

[7] Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. *MIS Quarterly*. 2003;27(3):425-478. doi:10.2307/30036540.

[8] Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*. 2012;36(1):157-178. doi:10.2307/41410412.

[9] Patil P, Tamilmani K, Rana NP, Raghavan V. Understanding consumer adoption of mobile payment in India: Extending Meta-UTAUT model with personal innovativeness, anxiety, trust, and grievance redressal. *International Journal of Information Management*. 2020;54:102144. doi:10.1016/j.ijinfomgt.2020.102144.

[10] Widyanto HA, Kusumawardani KA, Yohanes H. Safety first: extending UTAUT to better predict mobile payment adoption by incorporating perceived security, perceived risk and trust. *Journal of Science and Technology Policy Management*. 2022;13(4):952-973. doi:10.1108/JSTPM-03-2020-0058.

[11] Singh S. An integrated model combining ECM and UTAUT to explain users' post-adoption behaviour towards mobile payment systems. *Australasian Journal of Information Systems*. 2020;24. doi:10.3127/ajis.v24i0.2695.

[12] Srivastava S, Shunmugasundaram V. Mobile payment FinTech services adoption: comparing three models. *Management Decision*. 2025 (online first):1-29. doi:10.1108/MD-02-2025-0338.

[13] Motwani K, Choubey S, Saxena A, Patni I. Generation Z's mobile payment adoption: integrating UTAUT model and government initiative impacts. *Journal of Innovation and Entrepreneurship*. 2026;15(1). doi:10.1186/s13731-026-00658-1.

[14] Kharismawan R, Pramana E, Kristian Y. Factors Influencing Mobile Payment Adoption Utilizing Extended UTAUT with a Case Study: Public Transportation Trans Jatim Buses. *Jurnal Teknologi dan Manajemen Informatika*. 2025;11(1):69-80. doi:10.26905/jtmi.v11i1.15649.

[15] Ghosh A, Mishra R, Mishra A. Role of network externalities and trust facilitators in shaping mobile payment application continuance intentions: variations across stages of adoption. *Marketing Intelligence & Planning*. 2025;43(8):1553-1575. doi:10.1108/MIP-06-2024-0417.

[16] Ajao QM, Oludamilare O, Sadeeq L. Drivers of Mobile Payment Acceptance: The Impact of Network Externalities in Nigeria. *OALib*. 2023;10(06):1-19. doi:10.4236/oalib.1110240.

[17] Zheng W, Qiu H, Morrison AM. Applying a combination of SEM and fsQCA to predict tourist resource-saving behavioral intentions in rural tourism: An extension of the theory of planned behavior. *International Journal of Environmental Research and Public Health*. 2023;20(2):1349. doi:10.3390/ijerph20021349.

[18] Dieu HTM, Al Mamun A, Nguyen TLH, Naznen F. Cashless Vietnam: a study on intention and adoption of cashless payment. *Journal of Science and Technology Policy Management*. 2025;16(2):344-364. doi:10.1108/JSTPM-02-2022-0031.

[19] Chen M, Wang H, Liang Y, Zhang G. Net and configurational effects of determinants on cloud computing adoption by SMEs under cloud promotion policy using PLS-SEM and fsQCA. *Journal of Innovation & Knowledge*. 2023;8(3):100388. doi:10.1016/j.jik.2023.100388.

[20] Hussain M, Begum Q, Rahman MS, Imran A. Application of SEM and fsQCA to predict m-health adoption in the BoP market. *Digital Policy, Regulation and Governance*. 2023;25(6):545-567. doi:10.1108/DPRG-02-2023-0024.

[21] Niu Z, Liu W, Yao J, Xiao X, Wu J. A multidimensional analysis of potential suppliers' behavioral intentions in shared parking: Insights from SEM, NCA, and fsQCA. *Transport Policy*. 2025;174:103861. doi:10.1016/j.tranpol.2025.103861.

[22] Ragin CC. Redesigning Social Inquiry: Fuzzy Sets and Beyond. Chicago: *University of Chicago Press*; 2008.

[23] Fiss PC. Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*. 2011;54(2):393-420. doi:10.5465/amj.2011.60263120.

[24] Mattke J, Maier C, Weitzel T, Thatcher JB. Qualitative comparative analysis in the information systems discipline: a literature review and methodological recommendations. *Internet Research*. 2021;31(5):1493-1517. doi: 10.1108/INTR-09-2020-0529.

[25] Rasoolimanesh SM, Valaei N, Rezaei S. Guideline for application of fuzzy-set Qualitative Comparative Analysis (fsQCA) in tourism and hospitality studies. In: *Cutting Edge Research Methods in Hospitality and Tourism. Emerald Publishing Limited*; 2023:137-156. doi:10.1108/978-1-80455-063-220231009.

[26] Dul J. Necessary Condition Analysis (NCA): Logic and methodology of 'necessary but not sufficient' causality. *Organizational Research Methods*. 2016;19(1):10-52. doi:10.1177/1094428115584005.

[27] Dul J, van der Laan E, Kuik R. A statistical significance test for Necessary Condition Analysis. *Organizational Research Methods*. 2020;23(2):385-395. doi:10.1177/1094428118795272.

[28] Vis B, Dul J. Analyzing relationships of necessity not just in kind but also in degree: Complementing fsQCA with NCA. *Sociological Methods & Research*. 2018;47(4):872-899. doi:10.1177/0049124115626176.

[29] Solaimani S, Swaak L. Critical Success Factors in a multi-stage adoption of Artificial Intelligence: A Necessary Condition Analysis. *Journal of Engineering and Technology Management*. 2023;69:101760. doi:10.1016/j.jengtecman.2023.101760.

[30] Eccarius T, Chen CF. Examining trust as a critical factor for the adoption of electric vehicle sharing via necessary condition analysis. *Technological Forecasting and Social Change*. 2024;208:123681. doi:10.1016/j.techfore.2024.123681.

[31] Rasoolimanesh SM. Applying fsQCA, NCA and NConfA in tourism research. *Annals of Tourism Research*. 2026;116:104069. doi:10.1016/j.annals.2025.104069.

[32] Schneider CQ, Wagemann C. Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis. *Cambridge: Cambridge University Press*; 2012. doi:10.1017/CBO9781139004244.

[33] Greckhamer T, Furnari S, Fiss PC, Aguilera RV. Studying configurations with qualitative comparative analysis: Best practices in strategy and organization research. *Strategic Organization*. 2018;16(4):482-495. doi:10.1177/1476127018786487.

[34] Dul J. Conducting Necessary Condition Analysis for Business and Management Students. London: *SAGE Publications*; 2020.

[35] Dul J, Hauff S, Bouncken RB. Necessary condition analysis (NCA): review of research topics and guidelines for good practice. *Review of Managerial Science*. 2023;17(2):683-714. doi:10.1007/s11846-023-00628-x.

[36] Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. 1979;6(2):65-70.

[37] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*. 1995;57(1):289-300. doi:10.1111/j.2517-6161.1995.tb02031.x.

[38] Liu Y, Meng Q. How online reviews of sports tourism affect tourists' value co-creation behavior: Findings from PLS-SEM, NCA, and fsQCA. *Journal of*

Hospitality and Tourism Management. 2024;60:360-372. doi:10.1016/j.jhtm.2024.07.005.

[39] Yamin MAY, Abdalatif OAA. Examining consumer behavior towards adoption of quick response code mobile payment systems: transforming mobile payment in the fintech industry. *Humanities and Social Sciences Communications*. 2024;11(1). doi:10.1057/s41599-024-03189-w.

[40] Cao W, Chen Y, Wang K. What determines low-carbon travel? A policy-oriented assessment using PLS-SEM and NCA in an extended TPB framework. *Transportation*. 2026. doi:10.1007/s11116-026-10734-w.