

JOURNAL OF ACADEMY OF BUSINESS AND ECONOMICS TM

EDITOR-In-CHIEF

Professor Cheick Wague, South Stockholm University, Sweden



A Publication of the
International Academy of Business and Economics®

EDITORIAL BOARD:

Dr. Claudia Araiza, Southern States University, U.S.A.
Dr. Marc F. Audi, Paris1 Sorbonne, France
Dr. Antonina Bauman, Emporia State University, U.S.A
Dr. Adina Elena Danuletiu, Economics and Business Administration Department,
"1 Decembrie 1918" University Alba Iulia, Romania
Dr. Najat M. Jumaan, Sanaa University, Yemen
Dr. Vilas Khandare, Shri Asaramji Bhandwaldar College, Deogaon, India
Dr. Jerome Kueh, Swee Hui Universiti Malaysia Sarawak, Malaysia
Dr. Mayowa G. Agboola, Covenant University, Nigeria
Dr. Sunil Kumar, Amity University Haryana, India
Dr. Norma Laquian Menez, Lyceum of the Philippines University, Philippines
Dr. Zuraidah Mohd-Sanusi, Universiti Teknologi MARA, Malaysia
Dr. Long Dinh Pham, Open University HCMC, Vietnam
Dr. Bradley J. Prince, Associate Dean Richards College of Business,
University of West Georgia, U.S.A.
Dr. Lawrence P. Shao, Dean College of Business, Slippery Rock University, U.S.A.
Dr. Gilson G. Silva, UFSC, Brazil
Dr. Veena Tewari, Majan College, Oman
Dr. Wen-Hsien Tsai, National Central University, Taiwan
Dr. Sut Sakchutchawarn. School of Business & Economics, State University of New York,
Plattsburgh, U.S.A.

JABE is a publication of the International Academy of Business and Economics.
JABE is trademark of the IABE. All rights reserved. ©2026 IABE. WWW.IABE.ORG

Disclaimer: IABE or its representatives are not responsible for any error(s), validity of data/conclusion(s) or copyright infringements in any article published in the journal. Author(s) is/are solely responsible for the entire contents of the paper published in the journal.

TABLE OF CONTENTS

Page

THE IMPACT OF WORKPLACE SPIRITUALITY, LEADERSHIP, AND WORK-LIFE BALANCE ON THE JOB PERFORMANCE OF LADY OFFICERS IN THE SRI LANKA AIR FORCE Anuradha Iddagoda, University of Sri Jayewardenepura, Sri Lanka Rebecca Abraham, Nova Southeastern University, U.S.A. Manoj Keppetipola, Postgraduate institute of Management, University of Sri Jayewardenepura, Sri Lanka Hiranya Dissanayake, Wayamba University of Sri Lanka Madhura Jayasinghe, University of Sri Jayewardenepura, Sri Lanka	5
BRIDGING CULTURES: MEDIATION AS A PATHWAY TO INTERNATIONAL STUDENT SUCCESS Amr Swid, The Bill Munday School of Business, St. Edward's University, Austin, U.S.A. Joshua Bienstock, School of Management, New York Institute of Technology, New York, U.S.A. Nico Weigandt, The Bill Munday School of Business, St. Edward's University, Austin, U.S.A.	19
COVID-19 POLICY BUNDLES AND REGISTERED NURSE EMPLOYMENT RECOVERY: STATE-LEVEL EVIDENCE FROM THE UNITED STATES Karen Chin, Davis College of Business and Technology, Jacksonville University, U.S.A. John Shaw, Davis College of Business and Technology, Jacksonville University, U.S.A. Caitlin Sockbeson, Davis College of Business and Technology, Jacksonville University, U.S.A. Maggie Foley, Davis College of Business and Technology, Jacksonville University, U.S.A.	32
HOW BANK SOCIAL MEDIA ACTIVITIES SHAPE BRAND EQUITY: EVIDENCE FROM CUSTOMERS OF COMMERCIAL BANKS IN GHANA Fidelis Quansah, University of Professional Studies Accra, Ghana Timothy Kwabla Zilevu, University of Professional Studies Accra, Ghana Nathan Kobina Austin, University of the Virgin Islands, U.S.A. Julia Tomety, University of Professional Studies Accra, Ghana	50
MODELLING NON-LINEAR MARKET DYNAMICS: TECHNICAL ANALYSIS AND DEEP LEARNING PRICE FORECASTING FOR INDIAN ENERGY PSUs Veeramani Anandhabalaji, SRM Institute of Science of Technology, Kattankulathur, India Manivannan Babu, Bharathidasan School of Management, Bharathidasan University, India M.S. Narassima, Great Lakes Institute of Management, Chennai, India Balasundram Maniam, Sam Houston State University, Huntsville, TX, U.S.A.	66

REALIZED KURTOSIS AND VOLATILITY FORECASTING BEYOND VIX-IMPLIED VARIANCE Megan Y. Sun, School of Business and Economics, University of Wisconsin – River Falls, U.S.A. Dawn Hukai, School of Business and Economics, University of Wisconsin – River Falls, U.S.A.	77
DYNAMIC CAPABILITIES AND RESOURCE RECONFIGURATION IN THE JAPANESE EV INDUSTRY: THE CASE OF TOYOTA AND HONDA Woonghee Lee, Hanyang University, Seoul, South Korea	93
DIGITAL SERVQUAL FRAMEWORK FOR EXAMINING CUSTOMER PERCEPTIONS OF DIGITAL SERVICES IN ONLINE ACCOMMODATION PLATFORMS Susita Carroll, Winston-Salem State University, North Carolina, U.S.A.	105

THE IMPACT OF WORKPLACE SPIRITUALITY, LEADERSHIP, AND WORK-LIFE BALANCE ON THE JOB PERFORMANCE OF LADY OFFICERS IN THE SRI LANKA AIR FORCE

Anuradha Iddagoda, University of Sri Jayewardenepura, Sri Lanka

Rebecca Abraham, Nova Southeastern University, U.S.A.

Manoj Keppetipola, Postgraduate institute of Management, University of Sri Jayewardenepura, Sri Lanka

Hiranya Dissanayake, Wayamba University of Sri Lanka

Madhura Jayasinghe, University of Sri Jayewardenepura, Sri Lanka

[dx.doi.org/10.18374/JABE-26-3.1](https://doi.org/10.18374/JABE-26-3.1)

ABSTRACT

The purpose of this study is to investigate the impact of workplace spirituality, leadership, and work-life balance on a sample of 73 Lady officers in the Sri Lanka Air Force. While research on military organizations in developing countries is sparse, it is exceedingly so when the sample consists exclusively of Lady officers, given their limited representation in a typical Air Force. The study assesses if a spiritual approach is effective in increasing the engagement of Lady officers with their jobs, which in turn, enhances job performance. Data was collected from 73 Lady officers in the Sri Lanka Air Force using scales for workplace spirituality, leadership, work-life balance, employee engagement, and job performance designed by the authors, in 2025. Structural Equations Modeling using the PLS-SEM software was employed to analyze the data. Workplace spirituality and work-life balance significantly influenced job performance with mediation by employee engagement. Leadership, in contrast, failed to have a significant impact on job performance in the mediated model. Spirituality-based approaches that promote workplace spirituality and work-life balance may be successfully employed by the military to engage Lady flying officers to motivate them to improve job performance.

Keywords: workplace spirituality, leadership, work-life balance, employee engagement, job performance, military, Sri Lanka, Lady military officers

1. INTRODUCTION

The Sri Lanka Air Force consists of Male Officers, Lady Officers, Airmen and Airwomen. Over the past decade, an increasing number of women have joined the Air Force, being well represented in Lady Officer and Airwomen ranks. Although the literature has documented the human resource management variables influencing the motivation and job performance of men in these positions (see Iddagoda et al., 2025, for a review), there is a paucity of research on their female counterparts. This study sets forth a spirituality-based approach that positively influences employee engagement and job performance for a sample of female Air Force officers in the Sri Lanka Air Force. Employee engagement defined as cognitive engagement (work is a central life interest), emotional employee engagement (the employee is emotionally connected to the firm), and behavioral employee engagement (the employee engages in prosocial behaviors, such as reduced absenteeism and turnover) (Iddagoda et al., 2016). Results may be generalized to other developing countries.

The first contribution of this study lies in its choice of predictor variables. Workplace spirituality emerges from within the soul, transcending survival instincts (Garcia-Zamor, 2003). It is the collection of values, mores, and beliefs that nourish the soul in the workplace by finding deep meaning in work, social connectedness, and adherence to higher-level goals. This study views workplace spirituality as self-identification with the task to the extent that an emotional connection is forged between the self and the task that motivates employee engagement which elevates job performance. The second predictor is leadership. Intuitively, leadership is needed to channel the diverse motivations, desires, and predispositions of employees to joy and excitement with tasks. The third predictor is work-life balance, or maintaining a balance between personal and professional duties (Lott and Abendroth, 2023). Given that the entire career

span of Lady officers in the Sri Lanka Air Force is limited to 15-20 years, there is considerable pressure on Lady officers to excel at work in order to rise to high positions in the military. This quest for upward mobility coupled with the demands of children's schedules, responsibilities to extended family, and responsibilities to the community can escalate stress. The ability to achieve inner harmony between these competing goals may leave the employee with more time to engage in work and adopt strategies to improve job performance.

The second contribution of the study is the expansion of our knowledge of human resource management in military organizations. Military organizations are hierarchical, with a strict chain of command. It is worthwhile to examine if true identification with tasks and engagement with work can be achieved in such environments through self-motivation rather than explicit directives. Can an enriched inner self experience joy and meaningfulness in work in an environment that is less conducive to such self-expression ?

The third contribution of this study is the examination of workplace spirituality or work-life balance from the unique perspective of Lady officers in the Sri Lanka Air Force. We find no other examination in the literature of measures of these variables among Lady officers in a developing country. Sri Lanka is a deeply religious country. With a largely Buddhist population and other religions such as Hinduism, Christianity and Islam, Sri Lankans are heavily involved in visiting roadside shrines, chanting scriptures into the night, and making pilgrimages to domestic and international locations. It is likely that spirituality from religions forms the basis of workplace spirituality's tenet of the inner connection between the soul and work. The practice of communal gatherings in choirs, chanting of scriptures, and community prayers may lead to workplace spirituality's characteristic of social connectedness. The congruence between religious values, and social expressions of faith may lead to inner peace that leads to joy in engaging in work, and in turn, superior performance.

A fourth contribution of the study lies in measuring work-life balance in a developing country. In these environments, Females have many social responsibilities to extended family, religious services, and community organizations. As pilots in the Air Force, Lady officers undergo rigorous flight training. It is useful to measure if the reduction of stress between competing work and family responsibilities can enable employees to concentrate on work by becoming more engaged, leading to improved performance.

2. REVIEW OF LITERATURE

2.1 Conceptualization of Workplace Spirituality

Ashmos and Duchon (2000) created a seminal definition of workplace spirituality as "the recognition that employees have an inner life that nourishes and is nourished by meaningful work that takes place in the context of community"(Ashmos and Duchon, 2000: 137). Beyer (1999) set forth a simpler definition of humans as meaning-seeking animals who find meaning in work, and find meaning in relationships with social groups in the workplace.

Intuitively, workplace spirituality is the reaching out of the inner spirit of the employee to the task, to peers, and to the more abstract entity of the organization. This reaching out of the spirit enables the employee to find joy and ecstasy in work, connect with colleagues and supervisors through interpersonal communication, and pursue high-level goals with long-term benefits. Iddagoda et al. (2022) found direct organizational outcomes of these attributes. Joy and ecstasy in work was manifested as finding deep meaning in work (Wrzesniewski, 2003). Job tasks were viewed as exciting and enriching to the soul to the extent that they increased job satisfaction and organizational commitment.

Interpersonal interaction with peers combined with self-efficacy (confidence to complete job tasks) significantly explained job performance. We conjecture that team building and empathy for peers from social connectedness could lead to successful workplace collaborations, and effective conflict resolution. These outcomes may improve job performance. The third dimension of workplace spirituality is adherence to higher-level goals. Such goals enrich the inner self of the organization for the furthering of morality, empathy, respect, and compassion, so that the organization espouses morally upright positions.

2.2 Conceptualization of Leadership

Heaton et al. (2004) posited that “inside-out” leadership whereby there is a transformation of spirituality within the individual too the organizational context outside of them. This consists of leaders as facilitators of conflict resolution, assisting employees in overcoming fatigue, and creating an organizational culture that motivates employees to achieve performance goals. Further, leaders must support ethical standards, rewarding employees for “doing the right thing” (Jurkiewicz and Giacalone, 2004). Marques et al. (2005) set forth that the leadership required to achieve an enrichment of the inner self must support honesty, kindness, and courage in subordinates. These values must infuse the psyche of employees, thereby facilitating the achievement of inner peace.

Leadership must be empathetic toward employees’ family responsibilities. By enacting family-friendly programs, such as onsite childcare, providing staff to assist with errands, cooking, and cleaning, and flextime to enable employees to attend family events. Leaders reduce family stressors that can disrupt workers’ dedication to their work. Psychological capital, or inner peace at the individual level must be enhanced, as must social capital in group interaction (Garg and Sharma, 2025).

2.3 Conceptualization of Work-life Balance

There is a lack of literature on work-life balance from the perspective of the military. This study sets forth that Lady Air Force officers in patriarchal societies may work diligently to master the technical aspects of flight in order to prove their merit in comparison with their male counterparts. The shortened duration of their careers places more pressure on them to master flight training expeditiously. Yet, their membership in a patriarchal society places unusual family responsibilities upon them. They must care for elderly relatives, participate in religious festivals, join community groups, and attend weddings, funerals, and other family functions. These additional responsibilities are sufficiently time-consuming to reduce time spent at work. Maintaining a balance between the job stressors and family stressors may significantly enhance Lady officers’ inner peace, as they equitably fulfill both work responsibilities and family responsibilities, neither of which will be neglected.

3. HYPOTHESES DEVELOPMENT

3.1 Workplace Spirituality→Employee Engagement→Job Performance

Workplace Spirituality has a religious basis. In an overwhelmingly Buddhist nation, such as Sri Lanka, Buddhist values based on compassion, preservation of life, consistency, and self-discipline, are paramount for followers. These values align with the Iddagoda et al. (2016) conceptualization of employee engagement. Iddagoda et al. (2016) viewed employee engagement as the composite of cognitive engagement (the organization is the central life interest), emotional employee engagement (there is an emotional connection between the employee and job tasks), and behavioral employee engagement (prosocial behaviors such as reduced absenteeism, reduced turnover, diligent task completion, and speaking well of the organization to external entities. Compassion and preservation of life lead employees to view the organization as preserve of community life, and as provider for their families. Rewards, pay-for-performance incentives, and benefits are viewed as manifestations of compassion from the organization to the employee. Employees may respond by becoming emotionally attached to the organization through emotional employee engagement. Alternatively, they become so intensely loyal to the organization that they view it as a central life interest or display cognitive employee engagement. Consistency and self-discipline are achieved by Buddhists through all-night chanting of scriptures and regular participation in pilgrimages and visits to shrines. This consistency and self-discipline lead to the attributes of behavioral employee engagement, including reduced absenteeism, reduced turnover, diligent task completion, and praise of the organization.

Stewart and Brown’s (2011) conceptualization perceived job performance as consisting of the two dimensions of task performance and organizational citizenship. Increased employee engagement from workplace spirituality may contribute to task performance. Employees who view the organization as central

life interest are likely to prioritize work over other pursuits, leading to enhance job performance. Employees who are emotionally attached to their jobs will seek emotional strength by diligently completing tasks. Finally, employees who are behaviorally engaged in their work demonstrate regular attendance, retention, eagerness to acquire job skills, and diligently apply themselves to task completion.

Hypothesis 1: Workplace spirituality increases employee engagement, which improves job performance among Sri Lanka Lady Air Force officers.

3.2 Leadership→Employee Engagement→Job Performance

In a recent study of environmental sustainability among male Sri Lanka Naval officers, Iddagoda et al. (in press) found that leadership that inspired employees was effective in engaging employees and improving their job performance. Leaders described epic sea battles, the exploits of military heroes, parades in dress uniform, and displays of sophisticated modern weaponry. Subordinates were captivated by these tales, registering awe at the sight of military might, and feeling an overwhelming sense of pride at the sight of dress uniforms and parades. These images engaged the rank-and-file personnel, leading them to strive at their job tasks, thereby elevating job performance. In essence, leadership must be inspirational, in that leaders must exhort subordinates to give of their best by creating visual images of exceptional performance. In this vein, leaders in the Sri Lanka Air Force must inspire Lady officers to accept challenges, pursue tough flying missions, acquire skills that are unique and contribute significantly to the military. Lady officers must be empowered to the extent that they view flying dangerous and difficult missions, and becoming skilled in modern weapon systems, missiles, and aerial defenses, as central life interest (cognitive employee engagement). Leaders must provide resources, training, and equipment for Lady officers to succeed to the extent that they become emotionally attached to the organization. Finally, leaders should advise and mentor Lady officers to maintain regular attendance and improve retention for the full 15-20 years of employment. Such cognitive engagement, emotional engagement, and behavioral engagement with job tasks improves task performance. Leaders must also inform Lady of the benefits of early retirement after 15-20 years of service. In the post-military period, women may obtain advanced degrees leading to second careers. As children will be independent, women may become socially empowered by joining community organizations, volunteering for charitable work, and supporting environmental programs. These pursuits will enrich the lives of the women psychologically, as they become confident, and engage in the civic debates of the times. Women must learn about the monetary benefits of the post-retirement period, including state pensions with lifelong health provisions, the Sri Lanka Ex-Servicemen's Association that provides legal advocacy and social support through women's groups. They must take advantage of the full-tuition coverage for higher education, and salaries for household staff provided by the military.

The other dimension of job performance in the Stewart and Brown (2011) conceptualization is organizational citizenship. Organ (1988) defined organizational citizenship as the selfless giving of one's time to new employees or peers. Examples among Lady Air Force officers include volunteering at military events, assisting new employees, or leading women's organizations. Engaged employees are likely to exhibit organizational citizenship. If they view the Air Force as central life interest, are emotionally engaged in flying, attend regularly, and build long careers, they are more likely to be committed to the Air Force, thereby stimulating them to volunteer at company events or assist new employees.

Hypothesis 2: Leadership increases employee engagement, which in turn, improves job performance among Sri Lanka Lady Air Force officers.

3.3 Work-Life Balance→Employee Engagement→Job Performance

Thibaut and Kelley (1959) proposed social exchange theory. Social exchange theory envisions an equitable exchange between the employer and the employee. The employee applies himself or herself diligently to job tasks, engages in team building by working cooperatively in groups, and provides positive communication about the organization to external groups. Social exchange theory may be applied to work-life balance among Lady officers in the Sri Lanka Air Force. As the first generation of Lady Air Force officers, the women experience the desire to master the technical aspects of flight to fulfill their expectations that

they perform as effectively as their male counterparts. Additionally, Lady officers contribute to high-performance work teams tasked with special projects by participating in training and mentoring programs.

In exchange, the military organization is expected to create a work environment that assists Lady officers in effective task completion. One method of organizational assistance is providing the measures that promote work-life balance. Flextime enables Lady officers to substitute work hours for family time by permitting them to attend children's school events, accompany family members to doctors' appointments, and care for elderly parents. Onsite childcare at subsidized rates reduces the high cost of childcare. The provision of staff members to assist families who move to new military bases in finding schools, locating grocery stores, locating physicians, and moving into new properties, is yet another form organizational assistance to Lady officers. Lady officers are likely to view the provision of these amenities as an equitable exchange for their labor that reduces the conflicting time demands of work responsibilities and family responsibilities. Work-life balance may be achieved.

This study posits that the achievement of work-life balance through equitable social exchange instills feelings of gratitude among Air Force Lady officers. This appreciation for the Air Force may attach the women emotionally to the military in a manifestation of emotional employee engagement. The positive emotions aroused may stimulate striving to learn the intricacies of flight and engage fully in mentoring and training programs. The additional task engagement may result in superior task performance.

Hypothesis 3: Work-life balance increases employee engagement and job performance among Lady officers in the Sri Lanka Air Force.

4. METHODOLOGY

4.1 Data Collection

A member of the research team holds rank in the Sri Lanka Air Force. He approached the Air Force to secure permission to distribute the instrument to Lady officers. Upon securing permission, the research team administered the questionnaire to the participants in three waves in order to avoid common method variance.

A total of 73 Lady officers participated in the study. The sample consisted of 10 Pilot Officers, 23 Flying Officer 32 flight lieutenants, 8 squadron leaders. The sample was college-educated, with all participants having earned undergraduate degrees from Sir John Kotelawela Defence University, in spite of their entrant as a direct officer cadet to Sri Lanka Air Force Academy or Officer cadet to Sir John Kotelawela Defence University. As expected, tenure was commensurate with rank with pilot officers having less than 5 years of experience, flying officers having 5-8 years of experience, flight lieutenants having 9-11 years of experience, and squadron leaders having 12-15 years of experience.

4.2 The Survey Instrument

The authors created scales which were contained in the survey instrument administered to the participants (see Appendix). Workplace Spirituality was measured by a 22-item scale with items such as, "I try to help my co-workers relieve their suffering," and "My spirit is energized as a result of my work." Work-life Balance was measured by an 8-item scale with items such as, "My spouse has almost no complaints against my responsibilities at home," and "My children admire my role as father/mother." Leadership was measured by a 9-item scale with items such as, "My leader is a team builder," and "My leader provides challenging and meaningful work with opportunities for career development." Employee Engagement was measured by a 12-item scale with items such as, "I think that the most important thing that happened to me is involvement in my work/job," and "I believe that I live, eat, and breathe with my work/job." Job Performance was measured by a 7-item scale including items such as, "I am volunteering for extra work when it is necessary," and "I am not a person who takes unauthorized breaks, intentionally works slowly, and wastes military resources."

Cronbach alpha values ranged from 0.816 for the Job Performance Scale to 0.917 for the Leadership Scale, with all scales exceeding the required threshold of 0.7 (Nunnally and Bernstein, 1994). Two composite reliability measures had values ranging from 0.851 to 0.948 in excess of the threshold of 0.70, indicating adequate reliability. Convergent validity values of 0.503 (Workplace Spirituality) to 0.571 (Leadership) exceeded the minimum acceptable value of 0.50 (Hair *et al.*, 2019). Discriminant validity was evaluated using the Heterotrait-Monotrait criterion (Henseler *et al.*, 2015). Values ranged from 0.351 (Leadership to Employee Engagement) to 0.825 (Workplace Spirituality to Employee Engagement) above the recommended cutoff of 0.80.

The quality of the model was assessed using the coefficient of determination (R^2), effect sizes (f^2), predictive relevance (Q^2), and prediction accuracy (Root Mean Squared Error and Mean Average Error). The results show that the model explained 58.0% of the variance in Employee Engagement and 29.8% of the variance in Employee Job Performance. Effect sizes (f^2) were computed to assess the relative contribution of each predictor. The results indicate that the Employee Engagement $\rightarrow$ Job Performance ($f^2 = .424$) and Workplace Spirituality $\rightarrow$ Employee Engagement ($f^2 = .542$) represent large effects, while Work-life Balance $\rightarrow$ Employee Engagement ($f^2 = .069$) indicates a small-to-moderate effect. The Stone-Geisser's Q^2 values for Employee Engagement ($Q^2 = .540$) and Job Performance ($Q^2 = .334$) were > 0 , supporting predictive relevance of the model (Hair *et al.*, 2019). Prediction errors indicate that Employee Engagement (Root Mean Squared Error = .691, Mean Average Error = .548) was predicted with higher accuracy compared to Job Performance (Root Mean Squared Error = .837, Mean Average Error = .629).

4.3 Data Analysis

A path analytic model was created with paths from Workplace Spirituality to Employee Engagement, Leadership to Employee Engagement, Work-life Balance to Employee Engagement, and Employee Engagement to Job Performance. The PLS-SEM software was employed to test the significance of path coefficients, both through a direct model of each of the variable pairs listed above, and a mediated model in which Employee Engagement mediated the relationship of each predictor, i.e. Workplace Spirituality, Leadership, and Work-life Balance with Job Performance. Multicollinearity was examined using the Variance Inflation Factor (VIF). VIF values were 1.00 for Employee Engagement to Job Performance, 1.199 for Leadership to Employee Engagement, 1.585 for Workplace Spirituality to Employee Engagement, and 1.497 for Work-life Balance to Employee Engagement. Since all values are below the cutoff of 5.0 (Hair *et al.*, 2019), multicollinearity is not a concern for this model.

5. RESULTS

5.1 Direct Model of Hypotheses Testing

Hypothesis 1 stated that workplace spirituality increased employee engagement which improved job performance. The direct model of hypotheses tests is presented in Table 1. Table 1 shows that Hypothesis 1 was supported as workplace spirituality significantly increased employee engagement (path coefficient = 0.601, $t = 7.864$, $p < .001$), with employee engagement significantly improving job performance (path coefficient = 0.545, $t = 8.925$, $p < .001$). Hypothesis 2 stated that leadership increased employee engagement which increased job performance. Table 1 shows that Hypothesis 2 was not supported as leadership failed to significantly increase employee engagement (path coefficient = 0.053, $t = 0.813$, $p > .05$) although employee engagement significantly positively influenced job performance (path coefficient = 0.545, $t = 8.925$, $p < .001$). Hypothesis 3 stated that work-life balance increased employee engagement which improved job performance. Table 1 shows that Hypothesis 3 was supported as work-life balance significantly increased employee engagement (path coefficient = 0.208, $t = 2.592$, $p < .01$) though to a lesser extent than workplace spirituality. As expected, employee engagement significantly improved job performance (path coefficient = 0.545, $t = 8.925$, $p < .001$).

TABLE 1
DIRECT MODEL OF PATH COEFFICIENTS

Path	Path Coefficient	Sample Mean	Sample Standard Deviation	T-statistic	P-value
Workplace Spirituality→Employee Engagement	0.601***	0.598	0.076	7.864	0.00
Leadership→Employee Engagement	0.053	0.072	0.065	0.813	0.416
Work-life Balance →Employee Engagement	0.208**	0.210	0.080	2.592	0.01
Employee Engagement→Job Performance	0.545***	0.564	0.061	8.925	0.00

*p<.05, **p<.01 ***p<.001

Source: Author-created

5.2 Indirect Model of Hypotheses Testing

A mediation analysis was conducted to determine if employee engagement mediated the effects of workplace spirituality, leadership, and work-life balance on job performance. Table 2 shows the results of mediation analyses. Results show that the indirect effect of workplace spirituality on job performance via employee engagement was significant (path coefficient = 0.328, $t = 5.889$, $p < .001$). Likewise, the indirect effect of work-life balance on job performance through employee engagement was significant (path coefficient = 0.113, $t = 2.379$, $p < .01$). These findings indicate that employee engagement acts as a key mechanism through which workplace spirituality and work-life balance improve job performance.

In contrast, leadership failed to exert a significant influence on job performance through employee engagement (path coefficient = 0.029, $t = 0.776$, $p > .05$). This result suggests that leadership may need to be redefined in a contingency framework for it to influence job performance through employee engagement.

TABLE 2
MEDIATION MODEL I

Path	Path Coefficients	Sample Mean	Sample Standard Deviation	T-statistic	P-value
Workplace Spirituality→Employee Engagement→Job Performance	0.328***	0.338	0.058	5.689	0.00
Leadership→Employee Engagement→Job Performance	0.029	0.041	0.037	0.776	0.437
Work-life Balance→Employee Engagement→Job Performance	0.113**	0.119	0.048	2.379	0.017

*p<.05, **p<.01, ***p<.001

Source Author-created

6. CONCLUSIONS

6.1 Discussion of Results

This paper assessed the predictors of job performance in a unique sample of Lady officers in the Sri Lanka Air Force. It provides an insight into the human resource management variables motivating employee engagement and job performance in a population that is rarely examined, i.e., Lady pilots. The entry of Lady into flight training is a recent phenomenon, particularly in Sri Lanka where careers of Lady officers are limited to 15-20 years. The short duration of careers imposes a significant burden upon the women to prove their capabilities, while earning the respect of male counterparts.

The principal findings of the study were that workplace spirituality increased employee engagement, which in turn, improved job performance. Separately, work-life balance increased employee engagement and job performance. Although all paths were significant in the workplace spirituality→employee engagement→job performance, and the work-life balance→employee engagement→job performance models, the path from workplace spirituality to employee engagement had a higher path coefficient than that from work-life balance to employee engagement. We offer an explanation based on workplace spirituality being based on core values, while work-life balance is a reducer of work stress.

Workplace spirituality in Sri Lanka emanates from the core Buddhist beliefs of compassion, respect for peers, consistency, and responsibility. Same way other religions such as Christianity, Hinduism and Islam has played a similar role in such religious believers. These core beliefs are the essence of the Lady officer's being, bestowing a sense of self-identity. They nourish the inner self, enriching the spirit and conferring benevolent feelings that are carried into the work environment. Compassion and respect for peers encourage team building. Consistency leads Lady officers to practice flight maneuvers repeatedly until mastery is achieved, and to faithfully follow the directions of flight instructors. Responsibility to the military strengthens patriotism. During peace time, Lady officers, same as their male counterparts are involved in training, nation building projects and projects to promote environmental sustainability. They plant trees, clear waste from monuments, restore mangrove swamps, and engage in seed bombing projects (reforestation projects). Responsibility to protecting the country from climate change as a core value strengthens the resolve of Lady officers participating in projects to promote environmental sustainability. As the values of workplace spirituality are manifested in the workplace, employees begin to view the job as essential to their self-image, or employees become cognitively emotionally engaged. As the values that define them are manifested in the military environment, employees become emotionally engaged to the organization. The congruence between personal beliefs and the values of the organization lead to cognitive employee engagement and emotional employee engagement, both of which stimulate extra effort and striving that support job performance.

Work-life balance, on the other hand, is a reducer of work stress. It is a convenience function, rather than a core value. The employee appreciates the efforts undertaken by the employer in relieving work duties so that the employee is capable of fulfilling family duties. However, a convenience function does not inspire, capture interest, or stir the soul. It does not become the central life interest for a cognitively engaged employee. It fails to confer emotional attachment to the organization. However, an employee who experiences work-life balance may speak well of the organization, reduce absenteeism, reduce turnover, and apply themselves diligently to task completion. This is behavioral employee engagement. The extra diligence to work, such as flight training, and serving on high-performance teams, along with regular attendance and positive psychology results in improved job performance.

6.2 Theoretical Implications

The values-based definition of workplace spirituality shows fit with the conceptualization of workplace spirituality as deep meaning in work, social connectedness, and adherence to higher-order goals (see Iddagoda et al., 2022, for a review). Employees who find deep meaning in work do so due to values of consistency and responsibility. Respect and compassion underlie social connectedness, or rich

interpersonal interactions. Flight officers who feel a sense of responsibility to the military will see beyond mundane goals to long-term goals that ensure the military's survival and growth.

This paper adds to Shamir's (1991) definition of workplace spirituality as the produce of congruence of work and self-concept, where self-concept is developed by the use of core values to create positive self-image. The work environment becomes an extension of the inner self with interactions and interpersonal connections to be cherished and nurtured. Spillover theory (Diener, 1984) suggests horizontal spillovers with finding deep meaning in work spilling over into capturing employee interest through increased cognitive engagement and emotional engagement. Both forms of engagement were shown to increase task performance. Iddagoda *et al.* (2022) viewed job satisfaction as an outcome of finding deep meaning in work. Job satisfaction is a more fleeting affective state. This study's finding of employee engagement as a mediator of workplace spirituality and job performance suggests that workplace spirituality has a more consequential effect than altering an affective state, in that engagement is an enduring emotion which encourages employees to elevate job performance.

The paper adds to the nascent literature on work-life balance in developing countries. In a recent study, Garg and Sharma (2024) observed that psychological capital and social capital were enhanced by work-life balance. This study moves the analysis beyond the immediate effects of work-life balance to employee engagement and job performance. High levels of psychological capital and social capital are beneficial. However, they do not necessarily translate to active engagement with work, followed by superior job performance. The social exchange between the Air Force and the Lady officer provides an abiding resource that is mutually beneficial.

A related organizational outcome is organizational citizenship. It is possible that work-life balance may arouse feelings of appreciation among Lady officers, who increase behavioral employee engagement and eventually, organizational citizenship, i.e., work-life balance may stimulate Lady officers to volunteer at military events, and assist new employees. Future research should test the work-life balance→employee engagement→organizational citizenship relationship.

6.3 Managerial Implications

Training programs in the Sri Lanka Air Force should emphasize training in workplace spirituality. Human Resource departments should develop training materials that include role plays, skits, and games that set forth core values that connect the inner self to work responsibilities. The ensuing increase in workplace spirituality may increase cognitive employee engagement and emotional employee engagement.

The Sri Lanka Air Force should strengthen the conditions to promote work-life balance. Staff can be provided to assist new employees. Flextime may be implemented to grant more time to Lady officers to attend to their families. Appreciative employees are likely to reduce absenteeism, reduce turnover, and praise the military. They may recommend the Air Force as a family-friendly environment to their female friends.

6.4 Research Limitations and Recommendations for Future Research

Leadership failed to exhibit any predictive effect on employee engagement in both the direct model and the mediated model. Perhaps, the Leadership Scale needs to be revised to more closely reflect the value of Lady officers in the Sri Lanka Air Force.

No moderators were tested in this study. Religiosity may be a moderator of the workplace spirituality-employee engagement relationship. Religiosity is the extent to which an employee practices his or her religion. It may be hypothesized that employees with greater religiosity are likely to have higher levels of workplace spirituality, so that the two variables together will significantly increase employee engagement. Likewise, gratitude may be tested as a moderator of the work-life balance-employee engagement relationship. Employees who are grateful for the amenities provided by the military to assist them in fulfilling family responsibilities may demonstrate greater behavioral employee engagement, which, in turn, improves job performance.

Future studies may use a larger sample of Lady officers as more women join the Air Force, over time. The current sample size of 73 women can thus be increased, which is not possible at this time.

REFERENCES:

- Ashmos, Donde, and Duchon, Dennis. Spirituality at work: A conceptualization and measure, *Journal of Management Inquiry*, Volume 14, Number 1, Pages 20-39, 2000.
- Beyer, Janice M. Culture, meaning, and belonging at work, Paper presented at the 1999 *Chicago Academy of Management* meeting, Chicago, USA, 1999.
- Diener, Ed. Subjective well-being, *Psychological Bulletin*, Volume 95, Number 3, Pages 542-575, 1984.
- Garcia-Zamor, Jean Claude. Workplace spirituality and organizational performance, *Public Administration Review*, Volume 63, Number 3, Pages 355-363, 2003.
- Garg, Naval, and Sharma, Nidhi. Does spirituality at work promote work-to-family enrichment among Indian employees ? *International Journal of Organizational Analysis*, Volume 33, Number 8, Pages 1809-1830, 2025.
- Hair, Joseph F., Black, William C., Babin, Barry. E., and Anderson, Ralph E. *Multivariate Data Analysis*. Pearson Education, New Jersey, 2019.
- Heaton, Dennis P., Schmidt-Wilk, Jane, and Travis, Frederick. Constructs, methods, and measures for researching spirituality in organizations, *Journal of Organizational Change Management*, Volume 17, number 1, Pages 62-82, 2004.
- Henseler, Jorg, Ringle, Christian M., and Sarstedt, Marko. A new criterion for assessing discriminant validity in variance-based structural equations modeling, *Journal of the Academy of Marketing Science*, Volume 43, Number 1, Pages 115-135, 2015.
- Iddagoda, Anuradha, Abraham, Rebecca, Keppetipola, Manoj, Dassanaike, Hasantha. and Dissanayake, Hiranya. *Green Human Resource Management in Developing Countries: An Organizational Perspective in Sri Lanka*, Springer Nature, New York, New York, 2025.
- Iddagoda, Anuradha, Abraham, Rebecca, and Chinta, Ravi. Do the dimensions of workplace spirituality influence organisational outcomes ? A path analytic model and moderated regression model," *International Journal of Work Organisation and Emotion*, Volume 13, Number 4, Pages 372-393, 2022.
- Iddagoda, Anuradha, Opatha, H.H.D.N.P., and Gunawardana, Kennedy D. Towards a conceptualization and operationalization of the construct of employee engagement, *International Business Research*, Volume 9, Number 2, Pages 85-98, 2016.
- Jurkiewicz, Carole I., and Giacalone, Robert A. A values framework for measuring the impact of workplace spirituality on organizational performance, *Journal of Business Ethics*, Volume 49, Number 2, Pages 465-480, 2004.
- Lott, Yvonne, and Abendroth, Anja-Kristin. Affective commitment, home-based working and the blurring of work-hoe boundaries: Evidence from Germany, *New Technology, Work and Employment*, Volume 38, Number 1, Pages 52-102, 2023.

Marques, Joan, Dhiman, Satinder, and King, Richard. Spirituality in the workplace: Developing an integral model and comprehensive definition, *The Journal of American Academy of Business*, Volume 7, Number 1, Pages 81-91, 2005.

Nunally, Jum, and Bernstein, Ira. *Psychometric Theory* (3rd. ed), McGraw-Hill, New York, 1994.

Organ, Dennis. W. *Organizational Citizenship Behavior: The Good Soldier Syndrome*. D. C. Heath and Company, Lexington, MA, 1988.

Shamir, Boas. Meaning, self, and motivation in organizations, *Organization Studies*, Volume 12, Number 3, Pages 405-424, 1991.

Stewart, Gordon. L., and Brown, Kenneth G. *Human Resource Management: Linking Strategy to Practice*. Wiley, Somerset, N. J., 2011

Thibaut, John W., and Kelley, Harold H. *The Social Psychology of Groups*. Wiley, Somerset, NJ, 1959.

Wrzesniewski, Amy Finding positive meaning in work, In Cameron, K. S., Dutton, J. E., and Quinn, R. E. (Eds.) *Positive Organizational Scholarship*, pp. 296-306, Berrett-Koehler, San Francisco, 2003.

Appendix

Scale for Workplace Spirituality

Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1. I can easily put myself in other people's shoes.	☐	☐	☐	☐	☐
2. I am aware of and sympathize with others.	☐	☐	☐	☐	☐
3. I try to help my co-workers relieve their suffering.	☐	☐	☐	☐	☐
4. I am aware of my co-workers' needs.	☐	☐	☐	☐	☐
5. I do jobs or tasks automatically, without being aware of what I'm doing (reversed).	☐	☐	☐	☐	☐
6. I find myself working without paying attention (reversed).	☐	☐	☐	☐	☐
7. At work, I break or spill things because of carelessness, not paying attention, or thinking of something else (reversed).	☐	☐	☐	☐	☐
8. I rush through work activities without really being attentive to them (reversed).	☐	☐	☐	☐	☐
9. I go to the place on 'automatic pilot,' and then wonder why I went there (reversed).	☐	☐	☐	☐	☐
10. It seems I am working automatically without much awareness of what I'm doing (reversed).	☐	☐	☐	☐	☐
11. I experience joy in my work.	☐	☐	☐	☐	☐
12. I look forward to coming to work most days.	☐	☐	☐	☐	☐
13. I believe that others experience joy as a result of my work.	☐	☐	☐	☐	☐

14. My spirit is energized as a result of my work.	☐	☐	☐	☐	☐
15. I see a connection between my work and the larger social good of my community.	☐	☐	☐	☐	☐
16. I understand what gives my work personal meaning.	☐	☐	☐	☐	☐
17. The work I do is connected to what I think is important in life.	☐	☐	☐	☐	☐
18. At times, I experience an energy or vitality at work that is difficult to describe.	☐	☐	☐	☐	☐
19. I experience moments at work when everything is blissful.	☐	☐	☐	☐	☐
20. At times, I experience happiness at work.	☐	☐	☐	☐	☐
21. I have moments at work in which I have no sense of time or space.	☐	☐	☐	☐	☐
22. At moments, I experience complete joy and ecstasy in my work.	☐	☐	☐	☐	☐

Source: Author-created

Scale for Leadership

Please indicate how you feel about your experience in your job and your organization.

Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1. My leader is a team builder.	☐	☐	☐	☐	☐
2. My leader gives recognition for my achievements in the organization.	☐	☐	☐	☐	☐
3. My leader helps me to see and feel how they are contributing to the organization's success and time.	☐	☐	☐	☐	☐
4. My leader shows that he/she values me as an employee.	☐	☐	☐	☐	☐
5. My leader helps me create confidence in a company by being an example of high ethical and performance standards.	☐	☐	☐	☐	☐
6. My leader strives to maintain a company's reputation and demonstrate high ethical standards.	☐	☐	☐	☐	☐
7. My leader communicates a clear organizational vision.	☐	☐	☐	☐	☐
8. My leader clarifies his/her expectations of me as an employee and provides feedback on my functioning in the organization.	☐	☐	☐	☐	☐
9. My leader provides challenging and meaningful work with opportunities for career development.	☐	☐	☐	☐	☐

Scale for Employee Engagement

Please indicate how you feel about your experience in your job and your organization.

Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1. I think the most important thing that happened to me is involvement in my work/job.	☐	☐	☐	☐	☐
2. I believe the major satisfaction in my life comes from my work/job.	☐	☐	☐	☐	☐

3. I believe I live, eat and breathe with my work/job.	☐	☐	☐	☐	☐
4. When my Commanding Officer assigns a job/task I feel, I am really going to "get into" this job/task.	☐	☐	☐	☐	☐
5. I feel proud of the work I do.	☐	☐	☐	☐	☐
6. I am proud to introduce myself with my job title.	☐	☐	☐	☐	☐
7. I always arrive at work when I am expected to arrive.	☐	☐	☐	☐	☐
8. Generally, I am not a person of absenteeism.	☐	☐	☐	☐	☐
9. I exert high levels of effort to perform duties of my job.	☐	☐	☐	☐	☐
10. I speak positively about the organization when interacting with others.	☐	☐	☐	☐	☐
11. I have no intention to resign or retire immediately from my job.	☐	☐	☐	☐	☐
12. I strive towards achieving duties in the expected way by my organization.	☐	☐	☐	☐	☐

Source: Author-created

Scale for Work-life Balance

Please indicate your level of agreement or disagreement with regard to the following statements.

Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
27. My Commanding Officer considers me as an asset.					
28. My Commanding Officer has almost no complaints against me.					
29. My juniors/subordinates consider me as a role model for them to follow.					
30. My bother officers are happy in the way that I am interacting with them.					
31. Officers/Airmen who expect my services are satisfied about the way I serve them.					
32. My spouse has almost no complaints against my responsibilities at home.					
33. My children admire my role as a father/mother.					
34. My parents are really satisfied about the way that I am fulfilling my obligations and responsibilities towards them.					
35. I maintain a cordial relationship with my close relatives.					

Source: Author-created

AUTHOR PROFILES:

Anuradha Iddagoda, is a Senior Lecturer of the Management Science Unit, Department of Mathematics, Faculty of Applied Sciences, University of Sri Jayewardenepura Sri Lanka. She holds a PhD in Human Resource Management from the University of Sri Jayewardenepura Sri Lanka, Her research areas are

Human Resource Management, Green Human Resource Management, green orientation, employee engagement, leadership and work-life balance.

Rebecca Abraham is Professor of Finance at Nova Southeastern University. She obtained her doctorate in Finance from Alliant International University, San Diego, in 1989. She is the author of 142 publications including 5 books. She publishes in derivatives, corporate finance, and labor economics. She is Co Editor of the Journal of Economic and Administrative Sciences.

Air Vice Marshal Manoj Keppetipola, past Deputy Chief of Staff and Director General Administration of Air Force officer who holds a PhD from PIM - University of Sri Jayewardenepura. He also MBA from PIM, University of Jayewardenepura, having two Masters in Defence Studies from University of Kelaniya, and Masters in National Security and War Studies from National Defence University, Pakistan. He has published more than 20 academic journal articles including indexed journals and co-authored two Palgrave McMillan books which highlights his dedication to scholarly excellence.

Hiranya Dissanayake is a dedicated educator and researcher with a proven commitment to advancing knowledge in Corporate Governance, Sustainability, and Strategic Management Accounting. She has consistently demonstrated excellence in academia and has made a lasting impact on her students, peers, and the broader research community. Her extensive publication record of 50 academic journal articles highlights her dedication to scholarly excellence.

Madhura Jayasinghe is Professor of Food Science in the Department of Food Science and Technology at University of Sri Jayewardenepura. He has published in the areas of nutrition and plant health, having experimented with new plants and seeds designed to increase food security. He had 10 indexed publications, which along with book chapters, conference proceedings, and theses reports make up his research portfolio.

BRIDGING CULTURES: MEDIATION AS A PATHWAY TO INTERNATIONAL STUDENT SUCCESS

Amr Swid, The Bill Munday School of Business, St. Edward's University, Austin, U.S.A.
Joshua Bienstock, School of Management, New York Institute of Technology, New York, U.S.A.
Nico Weigandt, The Bill Munday School of Business, St. Edward's University, Austin, U.S.A.

[dx.doi.org/10.18374/JABE-26-3.2](https://doi.org/10.18374/JABE-26-3.2)

ABSTRACT

This paper explores the growing need for universities to address cultural clashes faced by international students. Cultural clashes are common, and they generate friction between students and institutions. Mediation, a structured conflict resolution process led by a neutral facilitator, is presented here as a useful tool for navigating these cultural complexities. By emphasizing shared interests over opposing positions and separating people from the problem, mediation supports constructive communication and intercultural competence. This paper draws on recent studies to examine the role of mediation in promoting international student adjustment. Those studies underscore the role of mediators in facilitating cross-cultural communication and mitigating acculturation stress among international students. By equipping students with communication tools and fostering empathy, mediation offers a promising avenue for supporting international student success.

Keywords: *International students, cultural conflicts, intercultural competence, conflict resolution, acculturation stress*

1. INTRODUCTION

The number of international students enrolled at universities worldwide has grown steadily, and with that growth comes a sharper need to understand how cultural differences shape student adjustment. Cultural friction is not an exception in these settings; it is a baseline condition that students and institutions must learn to manage.

Mediation, defined here as a structured process in which a neutral third party facilitates communication between disputing parties, has surfaced in the higher-education literature as a tool well-suited to this terrain. By creating space for dialogue across cultural lines, mediation gives students from diverse backgrounds a structured way to work through disagreement and, in doing so, supports the sense of belonging that anchor successful adjustment.

Two principles drawn from the broader negotiation literature inform how mediation operates in cross-cultural settings: emphasizing shared interests over fixed positions, and separating the people involved from the problem being addressed.

Recent work applies these principles directly to the university context. Kim and Mailey (2018) document the role of mediators in facilitating cross-cultural communication and conflict resolution among students from diverse backgrounds, and Smith and Khawaja (2017) examine how mediation can address acculturation stress and the interpersonal strains it produces in educational environments.

International students often arrive without a developed playbook for navigating conflict in their host culture. Mediation supplies that playbook in two ways. First, it gives students concrete communication strategies for resolving disagreements without escalation. Second, by opening channels for dialogue, it counters the isolation that can stall adjustment when students feel unheard or excluded. The remainder of this paper examines how mediation, alongside formal orientation programs, supports international student adjustment and how cultural differences shape the relationship between the two.

2. LITERATURE REVIEW

2.1 Academic Adjustment

Academic adjustment refers to the degree to which a student meets the cognitive, behavioral, and motivational demands of university coursework, class participation, study habits, performance under evaluation, and engagement with faculty. For international students, this process is layered: academic norms vary across educational systems, and what counts as competent classroom behavior in one country may read as passive or overconfident in another. Byram and Jones (2002) note that standardized institutional approaches tend to overlook these variations, and Ward, Bochner, and Furnham (2001) frame academic adjustment as one strand of a broader sociocultural adaptation that unfolds alongside psychological adjustment to a new environment.

Two factors shape how quickly and successfully this adjustment occurs. The first is the cultural distance between the student's prior academic system and the host institution's expectations for students from collectivist or high-power-distance cultures, for example, may approach faculty interaction differently than peers from individualist, low-power-distance settings (Hofstede, 2011; Chirkov, Vansteenkiste, Tao, & Lynch, 2008). The second is the support infrastructure the host institution makes available, including orientation programs, advising, and conflict-resolution channels such as mediation. Where this infrastructure is culturally responsive, academic adjustment improves; where it is generic, students from non-dominant cultural backgrounds tend to disengage (Shaughnessy, 2004).

2.2 The Use of Mediation and Academic Adjustment

The relationship between international students' group integration, academic adjustment, and the use of mediation does not reduce to a single mechanism. It is shaped by the cultural backgrounds students bring, the resources for conflict their institution makes available, and the fit between the two. Kaushal and Kwantes (2006) argue that cultural conflict arises from individual differences in how people interpret and respond to social and personal situations. These differences surface in attitudes toward aggression, openness, and the relative weight given to academic versus athletic pursuits (Stroink, 2005). When these differences go unaddressed, they harden into the kinds of disputes that derail adjustment.

University mediation offices are positioned to defuse this dynamic. Griffin (2002) describes these offices as handling a wide range of issues for students, staff, and the broader campus community, including interpersonal disputes. He notes that they offer advice, strategies, and education in alternative dispute resolution to help parties reach amicable outcomes. The intake function matters as much as the resolution function: by giving students a recognized channel, mediation offices make conflict legible rather than letting it accumulate as background stress.

Orientation programs sit upstream of mediation in this picture. Jameson (1998) reports a positive association between orientation participation and indicators of student success, including retention, academic performance, and use of on-campus support services such as mediation. The implication is that orientation is not a standalone intervention but a gateway; students who go through it are more likely to know that mediation exists, what it does, and when to use it. For this gateway to function across a culturally heterogeneous student body, however, orientation itself must be designed with cultural responsiveness in mind. Kaushal and Kwantes (2006) point to flexible leadership and creative programming as preconditions for accommodating diverse cultural backgrounds and reducing misunderstanding during the transition into university life.

H1: International students' group integration and academic adjustment are positively related to the use of mediation.

2.3 Cultural Differences and Academic Adjustment

Standardized orientation programs assume a homogeneous audience and, in doing so, tend to overlook the cultural backgrounds international students bring (Byram & Jones, 2002). This matters because culture

shapes how students process information, interact with peers and authority figures, and adapt to unfamiliar environments (Ward, Bochner, & Furnham, 2001). A program design that emphasizes individual initiative, "introduce yourself, ask questions, network actively", reads as common sense to students from individualist backgrounds and as confusing or even rude to students from collectivist ones (Shaughnessy, 2004). The reverse is also true: heavy reliance on group icebreakers can disorient students whose academic culture rewards independent navigation.

Chirkov, Vansteenkiste, Tao, and Lynch (2008) extend this argument by showing that cultural values, such as collectivism and uncertainty avoidance, shape students' receptivity to a formal orientation in the first place. Students from collectivist cultures gravitate toward orientation activities built around group cohesion and peer support; students from cultures high in uncertainty avoidance prefer structured programming with explicit guidelines and predictable expectations. The same orientation event can therefore generate very different adjustment outcomes depending on whose cultural lens is reading it.

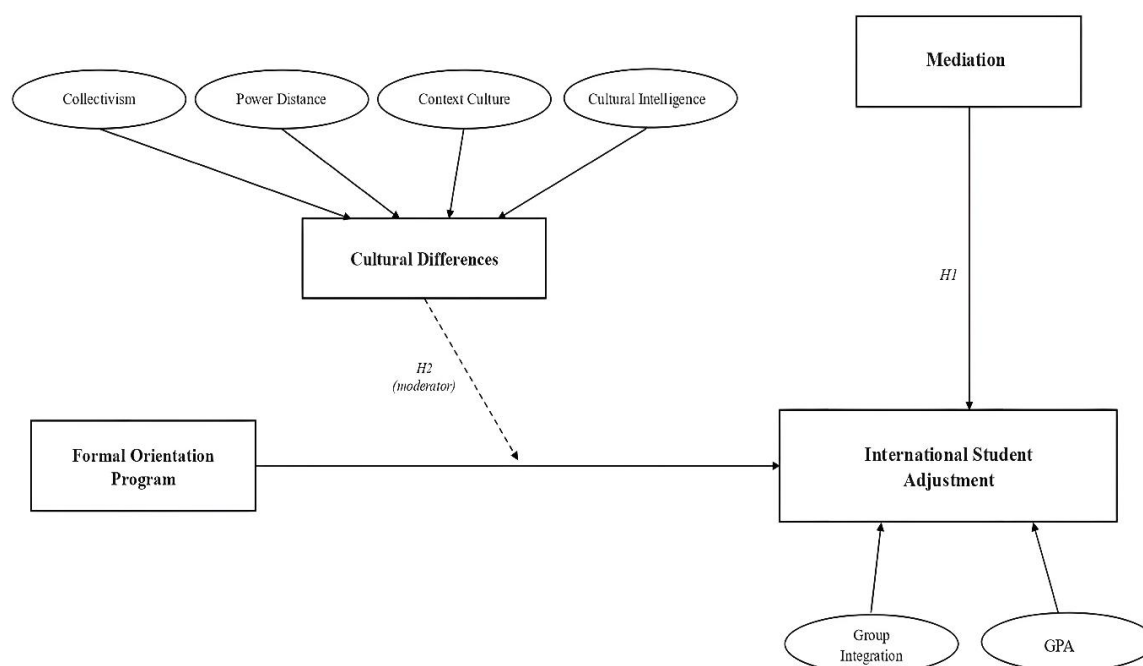
Hofstede's (2011) cultural dimensions framework adds a further layer through the power distance dimension. In high-power-distance cultures, hierarchical relationships are taken for granted, and international students from these backgrounds often defer to faculty and administrators and follow prescribed orientation norms without pushback. In low-power-distance cultures, students treat institutional authority as more horizontal and bring more autonomy and skepticism into the orientation process. Neither stance is intrinsically better, but a one-size-fits-all program will inadvertently reward one and penalize the other.

Context level, the degree to which communication relies on shared, implicit understanding versus explicit, direct statements, is a further moderator that orientation programs rarely account for. Students from high-context cultures read meaning from setting, relationships, and what is left unsaid; they expect faculty and peers to interpret indirect cues without forcing them into the open. When orientation materials and academic norms in the host country lean toward low-context explicit syllabi, direct office-hour requests, candid disagreement in seminars, high-context students can experience the environment as blunt or transactional, and their reluctance to ask questions directly can be misread as disengagement. The friction is not about a willingness to adjust; it is about a mismatch between the communication conventions students have internalized and those the institution defaults to (Hofstede, 2011; Ward, Bochner, & Furnham, 2001). Mediation is well-suited to surfacing this kind of mismatch precisely because it slows the conversation down enough for indirect signals to be heard.

Cultural intelligence, the capacity to function effectively across cultural contexts, sits alongside these dimensions as a student-level factor that conditions how cultural differences play out in practice. Students who arrive with stronger cultural-knowledge, motivational, strategic, and behavioral CQ are better equipped to read unfamiliar academic norms, adjust their behavior in real time, and recover from missteps without disengaging (Ang & Van Dyne, 2008; Chirkov et al., 2008). The implication for institutional design is that cultural difference is not a fixed obstacle to be removed but a variable to be worked with. Orientation programs that help students build cultural intelligence rather than simply transmitting institutional information produce more durable adjustment gains. This positions cultural intelligence as a plausible mediator between orientation participation and adjustment outcomes, a role the present study tests directly. Taken together, these studies establish cultural difference as a moderating variable in the relationship between formal orientation programs and international student adjustment. They also point to a research gap: most existing work treats orientation as a uniform treatment and culture as a static demographic, rather than testing how the interaction between the two produces adjustment outcomes. The present study addresses that gap. Universities can use these findings to design orientation programs that bridge cultural divides and equip international students with the support they need to succeed in their new academic setting.

H2: Cultural differences will moderate the relationship between formal orientation programs and international students' adjustment.

FIGURE 1: CONCEPTUAL MODEL



3. METHODS

3.1 Participants

The number of international students participating in this research study is 200, from heterogeneous countries of origin. The desired response would be 120 completed surveys from international students to ensure the accuracy and reliability of the research results. The participants will be randomly selected from three universities in New York State with a high percentage of registered international students. The students will all be incoming students who have not spent significant time in the US or attended any other schools in America. Respondents will be assigned an identification code and given an online address where they can complete the online survey through their customized code. All responses are confidential and will not be identified by the university.

3.2 Timing

The data will be collected longitudinally. A longitudinal study method will be used to detect the development of international students in US colleges. The first data collection from students will be during their orientation period. The second data collection will take place four months later, after the academic semester has ended. The advantage of a longitudinal study is that its findings go beyond a single moment in a student's life (Institute for Work & Health, 2009).

3.3 Measures

Power distance and collectivism. The perception of power distance and collectivism in the societies of origin of international students will be measured using the dimensions identified by the GLOBE Study (House, Hanges, Javidan, Dorfman, & Gupta, 2004). The survey focuses on the society in which the international

student lived before attending college in the United States. The survey has two types of questions, both based on a 1–7 Likert scale. For the first type, the student circles the number from 1 to 7 that best reflects their perceptions of their country. For the second type, the student circles the number from 1 to 7 that best reflects their level of agreement with the statement.

High vs. low context. To measure the cultural orientation and context of incoming international students, the Oddou and Derr (1999) survey, "Managing Internationally: A Personal Journey," will be used. The survey measures dimensions of culture, including Relationship, Space, Timing, and Communication. For this research, only the Relationship and Communication items will be used. The survey uses a 1–5 scale, where the student writes the number indicating their agreement with the statement, with 5 meaning strongly agree and 1 meaning strongly disagree.

An example statement to measure Relationship dimensions in culture is: "If I had some significant problems I needed help solving, I have any number of friends I could easily turn to for help." An example statement to measure Communication dimensions is: "If my boss or teacher were wrong, I would be more likely to tell her or him than to suggest there might be another answer simply."

Cultural intelligence. To measure incoming international students' cultural intelligence, the 20-item Four-Factor Cultural Intelligence Scale (CQS) developed by Ang and Van Dyne (2008) will be used. The CQS includes 20 items, divided into subcategories measuring CQ-Strategy, CQ-Knowledge, CQ-Motivation, and CQ-Behavior.

The survey uses a 1–7 scale, with 1 indicating strongly disagree and 7 indicating strongly agree. An example item to measure CQ-Strategy is: "I am conscious of the cultural knowledge I use when interacting with people with different cultural backgrounds." For CQ-Knowledge, an example item is: "I know the cultural values and religious beliefs of other cultures." An example item for CQ-Motivation is: "I enjoy interacting with people from different cultures." An example item for CQ-Behavior is: "I use pause and silence differently to suit different cross-cultural situations."

3.4 Orientation Program Participation

Participation in the university's formal orientation program was measured with a single item: "Did you attend the university's formal orientation program?" (1 = Yes, 0 = No). For those who attended, additional items assessed perceived helpfulness and content relevance on a 5-point Likert scale (e.g., "The orientation program helped me adjust to campus life," 1 = Strongly Disagree to 5 = Strongly Agree).

3.5 Mediation

Mediation will be assessed by the frequency with which international students have used the resource during their adjustment process. The question is: "Have you used any mediation resources during the past four months?" On a scale from 1 to 7, students will indicate the level and frequency of their use of mediation. A 1 represents "Never," a 4 represents "Sometimes," and a 7 represents "Frequently."

3.6 Adjustment Outcomes

Team/group integration. International students' team/group integration into their new university environments will be measured with a combination of four items from Morrison (1993) and three items from Chao, O'Leary-Kelly, Wolf, Klein, and Gardner (1994). Examples include: "My classmates and peers seem to accept me as one of them" and "I get along with people I work on group projects with very well" (Chao et al., 1994). Another example is: "I feel comfortable around my classmates and peers" (Chao et al., 1994). On a scale from 1 to 5, each participant indicates how much they agree or disagree with these items.

Grade Point Average. Grade Point Average is a measure of how well students are performing academically in college. In American universities, the GPA is based on a 4.0 scale, meaning that the closer the average is to 4.0, the better the student is performing in class.

4. ANALYSIS

4.1 Missing Value Analysis

Missing-value analysis is used to determine whether the dataset contains missing values. Missing values exceeding 5% in any dataset distort study results, so a missing-value analysis was conducted to determine whether missing values were present in the dataset. Only 1.82% of the data were missing; therefore, missing-value analysis was not necessary.

4.2 Normality

The multivariate kurtosis was less than 1.96 for all items, indicating that the data are normally distributed and can be used directly for further analysis.

4.3 Descriptive Statistics

Age. Students below 21 accounted for 41% of the sample, while students below 25 accounted for 85%, with a few graduate students (15%) representing older ages. The median age of the study participants was 21 years.

Citizenship. Most of the students were from China (41%), followed by India (22%). Other students were from European countries (7%), African countries (3%), and Arab countries (2%), totaling 14 countries.

Ethnic background. The majority were Asian (72%), followed by Caucasians, Africans, Arabs, and others.

4.4 Reliability

Reliability in statistics indicates the consistency of items within a construct. According to Table 1, the Cronbach's alpha in this study ranged from 0.556 to 0.936, indicating poor to superb consistency.

TABLE 1: SCALE RELIABILITY

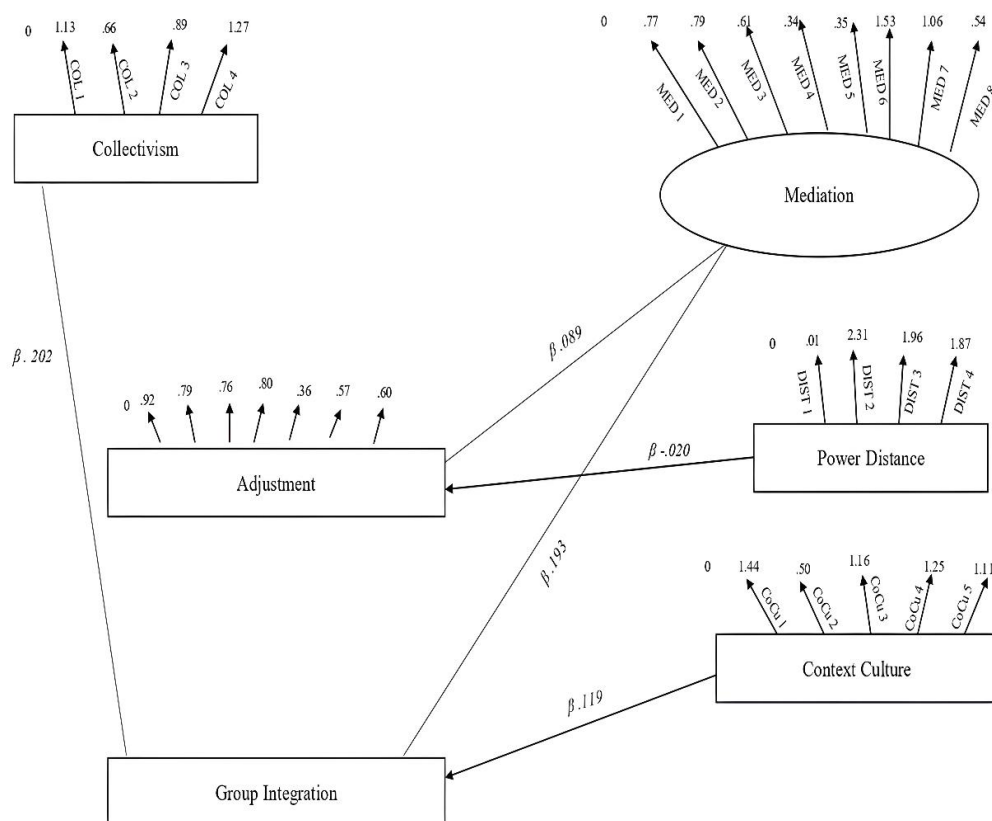
Particulars	Cronbach's alpha (α)
Academic adjustment	0.843
Mediation	0.912
Collectivism	0.645
Power distribution	0.556
Cultural intelligence	0.614
Low/high context culture	0.602
Group integration	0.936

4.5 Factor Analysis

When all items were forced to form a single factor, the factor analysis extracted 28 components, accounting for 70.375% of the variance. Since the factor loadings were greater than 0.4 (except for six items), the items with factor loadings below 0.4 were excluded from further analysis. The remaining items were retained for further analysis.

The path diagram for the preliminary model, used for confirmatory factor analysis, is shown in Figure 2.

FIGURE 2: PRELIMINARY MODEL



4.6 Model Fit Indices

The values of different indices are summarized in Table 2.

TABLE 2: PRELIMINARY MODEL FIT INDICES

	Model Fit	Desired Score
Chi-Square	9097.425	NA
Degrees of Freedom	4838	NA
CMIN/DF	1.88	≤ 2.00 for good fit; 2.00–5.00 for moderate fit
CFI	0.565	Close to or more than 0.90 for good fit
RMSEA	0.064	≤ 0.10 reflects good fit
NFI	0.387	A value close to 0.90 reflects a good fit

Table 2 shows that the chi-square value was 9097.425, the degrees of freedom were 4838, and the CMIN/DF was 1.88, indicating a good-fitting model. The CFI was 0.565, which falls short of the 0.90 threshold, and the RMSEA was 0.064, indicating a good fit. The NFI value of 0.387 is also below the 0.90 threshold.

4.7 Validity

Convergent validity. Principal component analysis was carried out to reduce large datasets into smaller, more meaningful sets of constructs. Each variable used in the analysis was measured by multi-item constructs, and factor analysis with varimax rotation was used to assess unidimensionality among the items. The constructs included in the confirmatory factor analysis had Cronbach's alphas greater than 0.7. Cronbach's alpha indicated good internal consistency among the constructs. All the items included in the analysis had factor loadings of 0.4 or greater. Convergent validity was also assessed using factor loadings, which had p-values less than 0.001. This test supported that the constructs had convergent validity.

Discriminant validity. Discriminant validity was indicated by the correlation matrix, in which the majority of the constructs had correlation coefficients below 0.85, and by the path analysis, in which the correlations among the latent constructs were below 1.

4.8 Measurement Model Fit Indices

The values of different indices for the measurement model are presented in Table 3.

TABLE 3: MEASUREMENT MODEL FIT INDICES

	Model Fit	Desired Score
Chi-Square	9296.938	NA
Degrees of Freedom	4838	NA
CMIN/DF	1.922	≤ 2.00 for good fit; 2.00–5.00 for moderate fit
CFI	0.544	Close to or more than 0.90 for good fit
RMSEA	0.065	≤ 0.10 reflects good fit
NFI	0.373	A value close to 0.90 reflects a good fit

Table 3 shows that the chi-square value was 9296.938, the degrees of freedom were 4838, and the CMIN/DF was 1.922, indicating a good-fitting model. The CFI was 0.544, and the RMSEA was 0.065. The NFI value of 0.373 is below the 0.90 threshold.

4.9 Hypotheses Testing

H1: International students' group integration and academic adjustment are positively related to the use of mediation.

The regression coefficient for the association between group integration and academic adjustment, controlling for mediation, was 0.193 ($p = 0.000$, < 0.05). Since the p-value is less than 0.05, we conclude that international students' group integration and academic adjustment are positively related to the use of mediation. Results are reported in Table 4.

TABLE 4: COVARIANCES: GROUP NUMBER 1 — DEFAULT MODEL

	Paths		β	S.E.	C.R.	P	Label
Adjustment	$\leftrightarrow$	Mediation	.089	.044	2.027	.043	
GroupInt	$\leftrightarrow$	Mediation	.193	.050	3.843	***	

H2: Cultural differences will moderate the relationship between formal orientation programs and the international students' adjustment.

Model structure. To test H2, a parallel-mediation model was estimated with formal orientation program participation as the predictor (X), international student adjustment as the outcome (Y), and two cultural variables as parallel mediators: context culture (M1) and cultural intelligence (M2). The model decomposes the total effect of orientation on adjustment into a direct path (Orient $\rightarrow$ Adjust) and two indirect paths routed through each mediator (Orient $\rightarrow$ ContCul $\rightarrow$ Adjust; Orient $\rightarrow$ CultInt $\rightarrow$ Adjust). Indirect effects were tested using bootstrapped 95% confidence intervals; an effect was considered statistically significant if the interval did not include zero. The sample size for this analysis was 120 international students. Results are reported in Table 5.

TABLE 5: DIRECT AND INDIRECT EFFECTS OF ORIENTATION ON ADJUSTMENT

Direct Effect of Orientation on Adjustment					
Effect	SE	t	p	LLCI	ULCI
.0302	.0173	1.7469	.0821	-.0039	.0642
Indirect Effect of Orientation on Adjustment					
	Effect	Boot SE	Boot LLCI	Boot ULCI	
Total	.0118	.0078	.0012	.0328	
ContCul	.0057	.0045	.0001	.0185	
CultInt	.0060	.0047	.0001	.0195	

To test H2, we ran a mediation analysis with orientation as the predictor, adjustment as the outcome, and two mediators sitting between them: context culture and cultural intelligence. Attending orientation gave students a small but real boost in both mediators (orientation to context culture: $\beta = 0.0408$, $p = .0231$; orientation to cultural intelligence: $\beta = 0.0429$, $p = .0376$).

From there, cultural intelligence pushed adjustment scores up in a statistically significant way ($\beta = 0.1410$, $p = .0293$), while context culture moved adjustment scores in the same direction but fell just short of the .05 cutoff ($\beta = 0.1404$, $p = .0585$), so we read it as marginal. Once both mediators were in the model, orientation on its own no longer moved adjustment to a meaningful degree ($\beta = 0.0302$, $p = .0821$).

What moved it was the indirect route. Bootstrapped confidence intervals showed a significant total indirect effect (0.0118, 95% CI [0.0012, 0.0328]), and each path separately held up too through context culture (0.0057, 95% CI [0.0001, 0.0185]) and through cultural intelligence (0.0060, 95% CI [0.0001, 0.0195]).

Read together, the numbers show that orientation does not improve adjustment on its own; it works by raising students' cultural awareness and cultural intelligence, and those gains carry the effect. That pattern aligns with H2's prediction: cultural factors are the channel through which orientation reaches adjustment.

5. FINDINGS AND DISCUSSION

This study set out to test whether cultural differences moderate the relationship between formal orientation programs and international student adjustment, and whether mediation contributes to adjustment alongside group integration. The results support both claims.

On the orientation side, the parallel-mediation analysis showed that orientation does not directly influence adjustment, but rather operates through cultural mechanisms, specifically context-culture awareness and cultural intelligence. This finding sharpens the existing literature on orientation programs in two ways. It corroborates earlier work documenting the limits of standardized orientation for diverse international student populations (Byram & Jones, 2002; Ward, Bochner, & Furnham, 2001), and it identifies the specific cultural channels through which orientation does generate adjustment gains. The implication is that cultural responsiveness is not a soft add-on to orientation programming; it is the active ingredient. Cultural factors such as collectivism, power distance, and uncertainty avoidance shape how students perceive and respond to orientation in the first place (Chirkov et al., 2008; Hofstede, 2011; Shaughnessy, 2004), so universities serious about adjustment outcomes need to design programs that account for these differences rather than treat them as noise.

On the mediation side, the H1 analysis confirmed that group integration and academic adjustment are positively associated with mediation use. This is consistent with the broader argument that communication, empathy, and cultural competence underwrite effective conflict resolution in cross-cultural settings (Griffin, 2002; Kaushal & Kwantes, 2006; Stroink, 2005). Mediation gives international students a structured channel to surface and resolve disagreements that, left unaddressed, accumulate into the kind of background stress that undermines adjustment.

The longitudinal design strengthens these interpretations. By collecting data at orientation and again four months later, the study captures adjustment as a process rather than a snapshot, allowing the indirect effects through cultural mediators to appear in the data.

Two boundary conditions are worth flagging. First, the direct effect of orientation on adjustment was non-significant in this sample; readers should not interpret this as evidence that orientation is unimportant, but as evidence that its mechanism is mediated. Second, the model fit indices (CFI and NFI) are well below the .90 threshold, suggesting that additional latent structure may be operating beyond what the present model captures. Future iterations should explore alternative model specifications and consider whether the construct definitions for power distance, low/high context, and cultural intelligence, several of which posted Cronbach's alphas below the conventional .70 threshold in this sample, would benefit from refined or extended item sets.

6. CONCLUSION

This study examined how formal orientation programs, cultural differences, and mediation jointly shape international students' adjustment to higher education. Two findings carry the paper. First, the effect of orientation on adjustment runs through cultural mediators, contextual culture, and cultural intelligence rather than operating directly, which reframes orientation programming as a vehicle for cultural-capacity building rather than information transfer. Second, mediation use is positively associated with both group integration and academic adjustment, which positions university mediation services as a meaningful complement to orientation rather than a niche dispute-resolution tool. Together, these findings shift the conversation about international student support from a one-size-fits-all orientation to a more layered, culturally responsive design.

6.1 Implications

The practical takeaway for university administrators is that orientation programs and mediation services should be designed and resourced as a connected system rather than as separate offerings. Three implications follow.

First, orientation content should be audited against the cultural composition of the incoming international student cohort. Programs that assume an individualist, low-context, low-power-distance student profile, which most standardized orientations leave students from other cultural backgrounds underserved. Tailoring communication style, the balance between group-versus-individual activities, and the explicitness of expectations is not a cosmetic adjustment; the mediation analysis in this study suggests that these design choices drive whether orientation translates into adjustment gains.

Second, mediation services should be introduced during orientation, not held in reserve until a conflict surfaces. Students who do not know mediation exists, or who associate it with formal grievance, will not access it during the early-semester window when adjustment friction is highest. Embedding a mediation overview into orientation lowers this access barrier.

Third, staff who deliver orientation and mediation services need training in cultural responsiveness, not as a one-off workshop but as an ongoing professional development track. The cultural-intelligence findings in this study apply to staff as well as students: a mediator's own CQ shapes how effectively they can hold space for cross-cultural disputes.

6.2 Future Research

Several directions extend this work. The first is on the staff development side: research is needed on which cultural competence training models for university staff most reliably translate into improved interactions with international students and a more inclusive campus climate. The training literature is uneven, and education-specific evidence is thinner than the corporate equivalent.

A second direction concerns the mediation process itself. The present study treats mediation use as a frequency variable; it does not unpack what makes a given mediation episode succeed or fail. Future work should examine mediator-level factors, such as cultural competence, communication style, and institutional positioning, and develop mediation protocols tailored to specific cultural conflict patterns rather than a generic dispute-resolution template.

Third, the cross-sectional and short- and long-term longitudinal evidence on mediation outcomes requires longer time horizons. The four-month follow-up window in this study captures early-semester adjustment but cannot speak to whether mediation engagement during the first year produces durable gains in academic performance, social integration, and overall well-being across a student's degree program.

Finally, the cultural-sensitivity and conflict-resolution principles examined here likely generalize beyond higher education to expatriate workforces, multicultural team settings, and other contexts where cultural difference and structured communication intersect. Comparative work across these settings would clarify which mediation principles are domain-specific and which transfer.

REFERENCES:

- Ang, S., & Van Dyne, L. (2008). *Handbook of cultural intelligence: Theory, measurement, and applications*. Armonk, NY: M.E. Sharpe.
- Byram, M., Gribkova, B., & Starkey, H. (2002). *Developing the intercultural dimension in language teaching: A practical introduction for teachers*. Council of Europe. <https://rm.coe.int/16802fc1b3>
- Chao, G. T., O'Leary-Kelly, A. M., Wolf, S., Klein, H. J., & Gardner, P. D. (1994). Organizational socialization: Its content and consequences. *Journal of Applied Psychology*, 79(5), 730–743. <https://doi.org/10.1037/0021-9010.79.5.730>
- Chirkov, V., Vansteenkiste, M., Tao, R., & Lynch, M. (2007). The role of self-determined motivation and goals for study abroad in the adaptation of international students. *International Journal of*

- Intercultural Relations, 31(2), 199–222. <https://doi.org/10.1016/j.ijintrel.2006.03.002> (Note: Your text cited 2008, but the publication date is 2007).
- Chao, G. T., O'Leary-Kelly, A. M., Wolf, S., Klein, H. J., & Gardner, P. D. (1994). Organizational socialization: Its content and consequences. *Journal of Applied Psychology*, 79, 730–743.
- Griffin, W. (2002). Academic dispute resolution and the role of the university ombuds. *Negotiation Journal*, 18(1), 73–89. <https://doi.org/10.1111/j.1571-9979.2002.tb00249.x>
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Sage Publications. (Note: If you specifically meant the 2011 article on dimensions, use: Hofstede, G. (2011). *Dimensionalizing cultures: The Hofstede model in context*. *Online Readings in Psychology and Culture*, 2(1). <https://doi.org/10.9707/2307-0919.1014>).
- House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (Eds.). (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Sage Publications.
- Institute for Work & Health. (2009). Cross-sectional vs. longitudinal studies. Retrieved from <https://www.iwh.on.ca/wrmb/cross-sectional-vs-longitudinal-studies>
- Jameson, D. A. (1998). The role of culture in international business: Building cultural sensitivity through orientation programs. *Business Communication Quarterly*, 61(4), 9–28.
- Kaushal, R., & Kwantes, C. T. (2006). The role of culture and personality in choice of conflict management strategy. *International Journal of Intercultural Relations*, 30(5), 579–603.
- Kim, E., & Mailey, M. (2018). Intercultural mediation in the college setting. *Negotiation Journal*, 34(4), 347–367.
- Morrison, E. W. (1993). Longitudinal study of the effects of information seeking on newcomer socialization. *Journal of Applied Psychology*, 78, 173–183.
- Oddou, G., & Derr, C. B. (1999). *Managing internationally: A personal journey*. Fort Worth, TX: The Dryden Press.
- Shaughnessy, M. F. (2004). Examining the association between academic factors and intercultural competence: Implications for international student intervention strategies. *College Student Journal*, 38(2), 225–235.
- Smith, R. J., & Khawaja, G. (2017). Acculturation stress and interpersonal mediation among international students: A conceptual model and measurement. *International Journal of Intercultural Relations*, 59, 8–21.
- Stroink, M. L. (2005). Personality, attitudes, and acculturation among international students. *Cross-Cultural Research*, 39(3), 263–292.
- Ward, C., Bochner, S., & Furnham, A. (2001). *The psychology of culture shock* (2nd ed.). Routledge.

AUTHOR PROFILE:

Amr Swid is Associate Professor of Management and MBA program Director at St. Edward's University. is an Associate Professor of Management and MBA Program Director at St. Edward's University. His research has been published in numerous leading international journals in the fields of management and organizational psychology. His scholarly work focuses on applied behavior analysis in global workplaces to improve organizational effectiveness and performance across both public and private sectors worldwide.

Joshua Bienstock is a Professor of Labor Relations and Business Law at New York Institute of Technology. He has been a practicing employment and labor lawyer for 30 years and a mediator for over 10 years. His research focuses on conflict resolution in the world of higher education and how different cultures contrast in their approach to conflict resolution.

Nico Weigandt earned a Bachelor of Science in Kinesiology at Texas A&M International in 2025. Currently, he is an MBA Candidate research assistant at St. Edward's University.

COVID-19 POLICY BUNDLES AND REGISTERED NURSE EMPLOYMENT RECOVERY: STATE-LEVEL EVIDENCE FROM THE UNITED STATES

Karen Chin, Davis College of Business and Technology, Jacksonville University, U.S.A.
John Shaw, Davis College of Business and Technology, Jacksonville University, U.S.A.
Caitlin Sockbeson, Davis College of Business and Technology, Jacksonville University, U.S.A.
Maggie Foley, Davis College of Business and Technology, Jacksonville University, U.S.A.

dx.doi.org/10.18374/JABE-26-3.3

ABSTRACT

The COVID-19 pandemic intensified pre-existing nursing shortages and contributed to uneven recovery in registered nurse (RN) employment across U.S. states. This study examines whether state COVID-19 policy environments in 2021 were associated with RN employment growth from 2021 to 2022. Using a cross-sectional sample of the 50 U.S. states and the District of Columbia (N = 51), this study models the percentage change in total RN employment from 2021 to 2022 as the dependent variable. To capture policy environments as bundles, individual policies are aggregated into three additive indices: Access and Affordability, Public Health Mandates, and Labor Mobility and Care Disruption. The third index is interpreted as a mixed labor-supply context because its components include both supports for labor-force attachment and family-care constraints. The models control for state-level economic and demographic conditions, along with a compensation-structure proxy. The primary GLM specifications, estimated with a Gaussian family, identity link, and Huber–White robust standard errors, produce directional patterns similar to those from OLS robustness checks, although the estimates are generally imprecise. The Access and Affordability index is positive across specifications, the Public Health Mandates index is negative and only marginally significant in the simpler specification, and the Labor Mobility and Care Disruption index is close to zero. Overall, the results provide limited evidence that broad state policy bundles alone explain cross-state variation in short-horizon RN employment recovery. From a managerial perspective, the findings support continuity-of-care and workforce-monitoring strategies during major disruptions, but they do not justify treating any policy bundle as a causal lever for RN employment.

Keywords: COVID-19 policy; telehealth; nursing workforce; employment recovery; state policy variation; mandates; labor mobility; health economics.

1. INTRODUCTION

The COVID-19 pandemic was a systemic shock that disrupted labor markets and health-care delivery simultaneously. In nursing, the shock arrived on top of pre-existing vulnerabilities: persistent shortage concerns, uneven geographic distribution, and high baseline burnout and turnover (Buchan & Aiken, 2008; World Health Organization, 2020). Because RN staffing is a binding constraint on service capacity, even modest changes in employment can translate into operational strain, greater reliance on temporary staffing, and higher labor costs. During 2020, health-care utilization and spending shifted sharply as elective and outpatient care was deferred and later rebounded unevenly (Karaca-Mandic et al., 2021). Nurse employment also changed during the early phases of the pandemic (Buerhaus et al., 2022).

A distinctive feature of the U.S. pandemic response was substantial state autonomy in policy design and implementation. States varied in policies affecting access to care, including telehealth coverage and reimbursement, prescription refill flexibility, and insurer cost-sharing waivers. They also varied in public health governance, including mask and vaccine mandates, and in policies tied to labor-force attachment through paid sick leave and school operations. This decentralization produced a heterogeneous policy landscape that can be linked to workforce outcomes.

Policy evaluation during a crisis is complicated because states rarely adopt interventions one at a time. Instead, interventions are implemented as bundles that jointly shape operational conditions. In a dataset

with only 51 jurisdictions, estimating separate coefficients for many correlated binary policies risks overspecification, unstable estimates, and multicollinearity. Bundle indices provide a parsimonious way to represent policy environments while aligning with a policy-mix perspective on crisis governance (Capano & Howlett, 2021; Foley, et. al., 2026).

This study asks whether state COVID-19 policy environments in 2021 are associated with RN employment recovery from 2021 to 2022. The policy environment is summarized by three additive indices: Access and Affordability, Public Health Mandates, and Labor Mobility and Care Disruption. The models control for state economic and demographic conditions, along with a compensation-structure proxy. The analysis is cross-sectional at the state level; because policy adoption is likely endogenous to outbreak severity, political constraints, baseline health-system capacity, and pre-existing workforce conditions, results are framed as descriptive associations rather than causal estimates throughout the paper.

Previewing the findings, the estimated coefficients are directionally consistent with the policy-bundle framework, but most are imprecisely estimated in this small cross-sectional sample. Access and continuity-of-care policies are positively associated with RN employment recovery, mandate-intensive environments are negatively associated, and the Labor Mobility and Care Disruption index remains close to zero over this short outcome window. These patterns are broadly similar across the primary Gaussian-identity GLM specifications with Huber–White robust standard errors and the OLS robustness checks. Rather than supporting strong claims about individual policy bundles, the results help refine future hypotheses about which policy environments may matter for workforce stabilization and motivate panel designs using richer workforce outcomes, including turnover, vacancies, hours worked, and migration.

2. LITERATURE REVIEW

A growing body of policy-mix research argues that policy evaluation should focus on combinations of instruments rather than single interventions in isolation (Capano & Howlett, 2021; Adolph et al., 2022). Cross-jurisdictional policy trackers document substantial variation in the timing and intensity of pandemic measures, supporting the use of indices to summarize policy environments and reduce dimensionality in small-sample state-level research (Hale et al., 2021; Cebula et al., 2025).

Access and affordability policies—including telehealth expansion, prescription refill flexibility, and cost-sharing relief—may support continuity of care and stabilize staffing needs. Telehealth scaled rapidly once regulatory and payment barriers were relaxed (Hollander & Carr, 2020; Wosik et al., 2020), and national evidence documents large increases in telemedicine use during the pandemic (Demeke et al., 2021; Patel et al., 2021). Emergency prescription refill flexibility may reduce avoidable disruptions by supporting medication adherence and enabling timely care seeking (Alexander & Qato, 2020; Dolan, 2020). Cost-sharing protections for COVID-19 treatment also evolved over time, with out-of-pocket costs for hospitalizations rising as some waivers expired (Chua et al., 2022).

Public health mandates aim to reduce transmission and protect health systems, but they can also introduce implementation frictions and political contestation. Reductions in mobility and consumer activity were driven partly by perceived infection risk rather than mandates alone, so mandate adoption may coincide with higher-risk environments and stronger local shocks (Chetty et al., 2020; Goolsbee & Syverson, 2021). In the health sector, evidence suggests that health care worker vaccine mandates did not produce large adverse employment effects on average (Wang et al., 2025).

Labor mobility and care-disruption policies can operate through offsetting labor-supply channels. Emergency paid sick leave can reduce presenteeism, facilitate isolation during illness, and support labor-force attachment; evidence links emergency sick leave provisions to reduced COVID-19 spread (Pichler et al., 2020). By contrast, school and childcare disruptions constrained labor supply and hours worked, especially among parents (Collins et al., 2021; Garcia & Cowan, 2022; Landivar et al., 2023). More broadly, the pandemic's uneven caregiving burden has been linked to differential labor-market outcomes (Goldin, 2022; Amuedo-Dorantes et al., 2023). For this reason, the Labor Mobility and Care Disruption index should

not be interpreted as a uniformly protective policy bundle. It captures both labor-attachment support and family-care constraints, so its net association with RN employment recovery may be weak or ambiguous. Cross-state labor mobility may amplify policy effects by enabling workers to reallocate toward higher-demand or better-compensated markets. Migration theory emphasizes push-pull forces and adjustment costs (Lee, 1966), while housing costs and spatial frictions can limit labor reallocation (Ganong & Shoag, 2017; Cebula & Clark, 2011).

Cost-of-living and compensation conditions may also shape nurse labor supply and retention. Higher housing costs can deter in-migration or encourage out-migration (Ganong & Shoag, 2017). Pay dispersion may also proxy for workplace disutility, as prior research links pay inequality to morale, cooperation, and performance (Cowherd & Levine, 1992; Bloom, 1999; Shaw et al., 2002).

2.1 Hypotheses

H1: Higher values of the 2021 Access and Affordability index are positively associated with RN employment recovery.

H2: Higher values of the 2021 Public Health Mandates index are negatively associated with RN employment recovery.

H3: Higher values of the 2021 Labor Mobility and Care Disruption index are associated with RN employment recovery, but the expected direction is ambiguous because paid sick leave may support labor-force attachment while school-related policies may reflect family-care constraints and community disruption. The relationship may therefore be weak or close to zero in a one-year net-employment framework.

H4: RN employment recovery is associated with structural state characteristics, including housing costs, compensation structure, unemployment, GDP per capita, Tax burden, and the share of the population aged 65 and older.

3. DATA AND METHODS

This study uses publicly available U.S. state-level data from 2021 and 2022. The unit of analysis is the state (50 states plus the District of Columbia; $N = 51$). RN employment data are drawn from the Bureau of Labor Statistics Occupational Employment and Wage Statistics (OEWS) state files for 2021 and 2022. Policy variables are compiled from the KFF COVID-19 State Data and Policy Actions Tracker and state public sources.

Controls are drawn from the BLS Local Area Unemployment Statistics, the American Community Survey, and the Federal Housing Finance Agency House Price Index. State and local Tax burden measures are taken from the Tax Foundation's state-level estimates.

3.1 Outcome Variable

The dependent variable is the state-level percentage change in total employed RNs, denoted $\% \Delta RN_i$. It represents the year-over-year change in RN employment between 2021 and 2022 and is calculated as:

$$\% \Delta RN_i = \frac{RN_{i,2022} - RN_{i,2021}}{RN_{i,2021}} \times 100$$

where $RN_{i,2021}$ and $RN_{i,2022}$ are total RN employment in state i in 2021 and 2022, respectively.

Using a relative rather than absolute change ensures comparability across states with different population sizes and workforce baselines. This measure aligns with prior macroeconomic analyses that use year-to-year employment growth rates to evaluate recovery from labor shocks (Coibion, Gorodnichenko, & Weber, 2020; Goolsbee & Syverson, 2021).

3.2 Variables and Methodology

To reduce model complexity and limit multicollinearity among closely related policy indicators, individual COVID-19 policies are grouped into three additive policy indices: Access and Affordability, Public Health Mandates, and Labor Mobility and Care Disruption. Each index is measured as a count of relevant binary policy indicators in 2021. A one-unit increase in an index indicates that a state adopted one additional policy within that category. This grouping strategy preserves interpretability in a 51-jurisdiction sample while avoiding an overspecified model with too many correlated policy variables.

Group 1: Access and Affordability. This index includes telehealth expansion, early prescription refill flexibility, and insurer cost-sharing waivers.

Group 2: Public Health Mandates. This index includes mask mandates and vaccine mandates, as coded in the dataset.

Group 3: Labor Mobility and Care Disruption. This index includes paid sick leave and school-related policies. This bundle combines policies that may affect labor supply in different directions. Paid sick leave may support labor-force attachment, while school-related disruptions may create childcare burdens and reduce work availability. Therefore, the coefficient on this index is interpreted as a net association across potentially offsetting mechanisms, rather than as the effect of a uniformly protective policy environment.

The models include controls for unemployment in 2021, GDP per capita, housing prices, Tax burden, the share of the population aged 65 and older, and an executives-to-RN pay ratio. Macroeconomic inflation is not included as a separate regressor because it is constant across states in this single-period cross-sectional design. Cross-state differences in local economic conditions and purchasing-power pressure are instead partially captured by GDP per capita and the Housing Price Index.

States did not enter the pandemic with the same nursing workforce conditions. Some states already had more nurses, greater hospital capacity, or more severe nursing shortages before COVID-19. Because consistent RN vacancy data are not available for all states, this paper treats pre-existing nursing workforce capacity as a limitation. To partially address this issue, the robustness checks include lagged RN employment.

The empirical strategy estimates the association between policy-group indices and RN employment growth using generalized linear models (GLM) with a Gaussian family and identity link as the baseline approach. Huber–White robust standard errors are reported throughout. Ordinary least squares (OLS) models with the same covariates are reported as robustness checks.

Four sequential models are estimated.

Model A includes only the Access and Affordability index:

$$\% \Delta RN_i = \beta_0 + \beta_1 G1_i + \gamma' X_i + \varepsilon_i \quad (1)$$

where $G1_i$ is the Access and Affordability index.

Model B includes only the Public Health Mandates index:

$$\% \Delta RN_i = \beta_0 + \beta_2 G2_i + \gamma' X_i + \varepsilon_i \quad (2)$$

where $G2_i$ is the Public Health Mandates index.

Model C includes only the Labor Mobility and Care Disruption index:

$$\% \Delta RN_i = \beta_0 + \beta_3 G3_i + \gamma' X_i + \varepsilon_i \quad (3)$$

where $G3_i$ is the Labor Mobility and Care Disruption index.

Model D includes all three policy indices simultaneously:

$$\% \Delta RN_i = \beta_0 + \beta_1 G1_i + \beta_2 G2_i + \beta_3 G3_i + \gamma' X_i + \varepsilon_i \quad (4)$$

In all models, X_i represents the vector of state-level control variables, and ε_i is the error term. This sequence first evaluates each policy bundle separately and then evaluates the three indices jointly in the full model. The same specification sequence is used in the robustness checks.

4. EMPIRICAL RESULTS

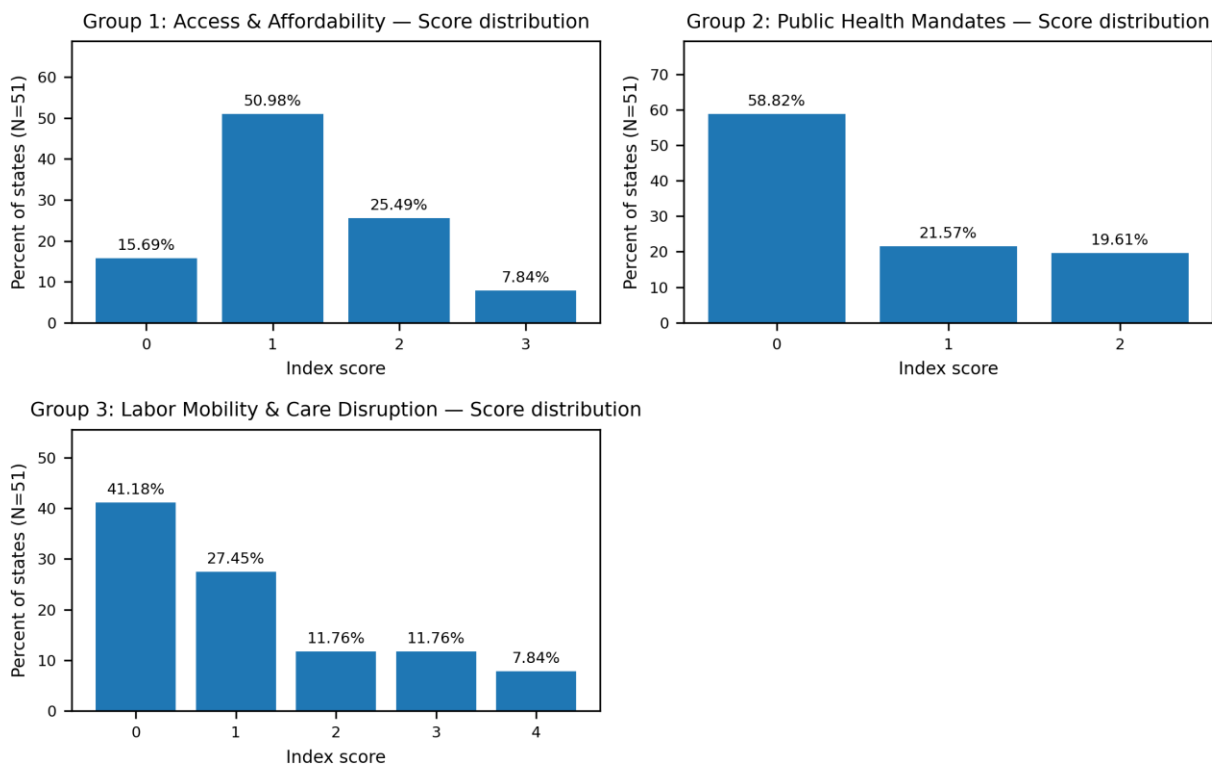
4.1 Descriptive Analysis

Table 1 reports the distribution of index scores for the three grouped COVID-19 policy domains across the 50 states and the District of Columbia. Scores represent the number of policies adopted within each domain in 2021. Group 1 (Access & Affordability) is concentrated at low-to-moderate adoption levels: a majority of states score 1 (26 states; 50.98%), while 8 states (15.69%) adopted none and only 4 states (7.84%) adopted all three measures. Group 2 (Public Health Mandates) is more polarized, with most states adopting neither mandate (30 states; 58.82%), and the remaining states split between adopting one mandate (11 states; 21.57%) or both mandates (10 states; 19.61%). Group 3 (Labor Mobility & Care Disruption) shows the widest spread across the scale: 21 states (41.18%) score 0, 14 states (27.45%) score 1, and a smaller subset of states adopted three or four policies (6 states, 11.76%, and 4 states, 7.84%, respectively). Taken together, the distributions indicate that access-oriented measures were broadly implemented, while mandate-based and labor mobility/care disruption policy environments were concentrated in a smaller set of states, generating meaningful cross-state variation for subsequent regression analysis.

Table 1. Policy-index score distributions (N = 51)

Score	Group 1: Access & Affordability	Group 2: Public Health Mandates	Group 3: Labor/Care
0	8 (15.69%)	30 (58.82%)	21 (41.18%)
1	26 (50.98%)	11 (21.57%)	14 (27.45%)
2	13 (25.49%)	10 (19.61%)	6 (11.76%)
3	4 (7.84%)	—	6 (11.76%)
4	—	—	4 (7.84%)
Total	51 (100%)	51 (100%)	51 (100%)

Note. “—” indicates that the score is not possible for that index given the number of component policies (Group 1 ranges 0–3; Group 2 ranges 0–2; Group 3 ranges 0–4). Percentages are shares of the full sample (N = 51).



Figures 1–3. Score distributions for the three policy-group indices: Group 1 (Access & Affordability), Group 2 (Public Health Mandates), and Group 3 (Labor Mobility & Care Disruption).

Table 2 reports descriptive statistics for the sample of 51 jurisdictions (50 states plus the District of Columbia). The dependent variable, RN employment recovery, averages 0.74 percentage points (SD 4.21) and ranges from -7.57 to 11.06, indicating substantial cross-state variation in the pace of recovery between 2021 and 2022. Among the key controls, the executives-to-RN pay ratio averages 2.65 (SD 0.40), unemployment averages 3.39 (SD 0.76), GDP per capita averages 64.83 thousand dollars (SD 10.00), housing price index averages 432.45 (SD 112.67), Tax burden averages 0.046 (SD 0.026), and the share of the population aged 65+ averages 0.226 (SD 0.024). Inflation is not listed in the descriptive table because it has no cross-jurisdictional variation in this single-period cross-sectional design.

Table 2. Descriptive Statistics (Overall)

Variable	n	mean	Standard deviation	min	max
RN employment recovery (%)	51	0.7385	4.2113	-7.5702	11.0561
pay_ratio	51	2.6549	0.3997	1.7741	3.5744
unemployment_rate	51	3.3941	0.7559	1.9000	5.2000
GDP per capita (\$1,000s)	51	64.8313	10.0036	47.0860	100.1530
housing_price	51	432.45	112.67	286.47	745.44
tax_rate	51	0.0461	0.0261	0.0000	0.0930
pct_65plus	51	0.2263	0.0238	0.1600	0.2700

4.2 Correlations

Table 3 reports the pairwise correlation matrix for the outcome, policy-group indices, and controls. RN employment recovery is only weakly correlated with the policy indices: it is positively associated with Group 1 (Access & Affordability) ($r = 0.2040$), slightly negatively associated with Group 2 (Mandates) ($r = -0.1146$), and essentially uncorrelated with Group 3 (Labor Mobility & Care Disruption) ($r = 0.0223$). The strongest correlations occur among the policy indices themselves, particularly between Group 2 and Group 3 ($r = 0.6398$), suggesting meaningful overlap between mandates and labor mobility/care disruption-support policy environments. Group 1 also covaries moderately with Group 3 ($r = 0.3301$) and with Group 2 ($r = 0.2157$).

Among controls, Group 2 is moderately correlated with unemployment ($r = 0.5636$), GDP per capita (\$1,000s) ($r = 0.5094$), and housing prices ($r = 0.3280$), and Group 3 shows similar positive correlations with unemployment ($r = 0.4141$), GDP per capita (\$1,000s) ($r = 0.4974$), and housing prices ($r = 0.3562$). Group 1 is positively correlated with GDP per capita (\$1,000s) ($r = 0.3123$) and housing prices ($r = 0.3239$) and negatively correlated with the Executives-to-RN pay ratio ($r = -0.2292$). Population age 65+ (%) is negatively correlated with GDP per capita (\$1,000s) ($r = -0.3264$) and housing prices ($r = -0.2289$). Overall, the correlation structure indicates limited bivariate association between the outcome and the policy indices, but notable inter-correlation across policy groups - especially between Groups 2 and 3 - highlighting the importance of multivariable models and collinearity diagnostics in the regression analysis.

Table 3. Correlation Matrix

Variable	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10
RN employment recovery (%)	1.0000									
G1: Access & Affordability	0.2040	1.0000								
G2: Public Health Mandates	-0.1146	0.2157	1.0000							
G3: Labor/Care	0.0223	0.3301	0.6398	1.0000						
Executives-to-RN pay ratio	-0.0670	-0.2292	0.1204	0.2463	1.0000					
Unemployment rate	0.0999	0.1959	0.5636	0.4141	-0.0012	1.0000				
GDP per capita (\$1,000s)	0.1198	0.3123	0.5094	0.4974	0.1330	0.1573	1.0000			
Housing price index	0.0501	0.3239	0.3280	0.3562	-0.0960	0.0004	0.5104	1.0000		
Tax rate	-0.1867	0.2013	0.3956	0.3343	-0.0090	-0.0143	0.0854	0.1867	1.0000	
Population age 65+ (%)	-0.0754	-0.1857	-0.0675	-0.1326	0.0046	-0.1600	-0.3264	-0.2289	0.0203	1.0000

4.3 Model Diagnostics: Multicollinearity, Residuals, and Influence (VIF, Residual Plots, Cook's Distance)

To assess multicollinearity among the policy indices and control variables, variance inflation factors (VIFs) were computed. As reported in Table 4, multicollinearity is modest: the mean VIF is 1.86, and the highest VIF is 3.21 for the Group 2 (Public Health Mandates) index. All other predictors range between 1.23 and 2.24 (tolerances ≥ 0.311). These values are well below conventional thresholds (e.g., 5 or 10), indicating that multicollinearity is unlikely to materially affect coefficient stability or inference in the multivariable models.

Table 4. Variance Inflation Factors (VIF)

Label	Variable	VIF	Tolerance
Group 1: Access & Affordability (0-3)	group1_access	1.40	0.713
Group 2: Public Health Mandates (0-2)	group2_mandates	3.21	0.311
Group 3: Labor/Care (0-4)	group3_context	2.24	0.446
GDP per capita (\$1,000s)	Gdp_k	2.16	0.462
unemployment_rate	unemployment_rate	2.11	0.474
housing_price	housing_price	1.59	0.631
tax_rate	tax_rate	1.53	0.652
pay_ratio	pay_ratio	1.28	0.781
pct_65plus	pct_65plus	1.23	0.813

Note. Mean VIF = 1.86.

Figures 4 and 5 evaluate the distribution of residuals from the baseline GLM specification (Gaussian family with identity link; Model D). The residual histogram is approximately symmetric and centered near zero, indicating no systematic deviation in the conditional mean specification. The Q–Q plot shows that residuals closely follow the theoretical normal reference line over the central range, suggesting that model fit is adequate for the bulk of observations. Deviations from normality are concentrated in the tails, particularly the upper tail, indicating a small number of jurisdictions with larger-than-predicted positive residuals in RN employment recovery. Given the relatively small cross-sectional sample ($N = 51$), mild tail deviations are expected in state-level data and do not necessarily indicate misspecification.

To ensure robustness to heteroskedasticity and tail behavior, all GLM estimates are reported with Huber–White robust standard errors. In addition, influence diagnostics are conducted to assess whether a small number of jurisdictions disproportionately affect results.

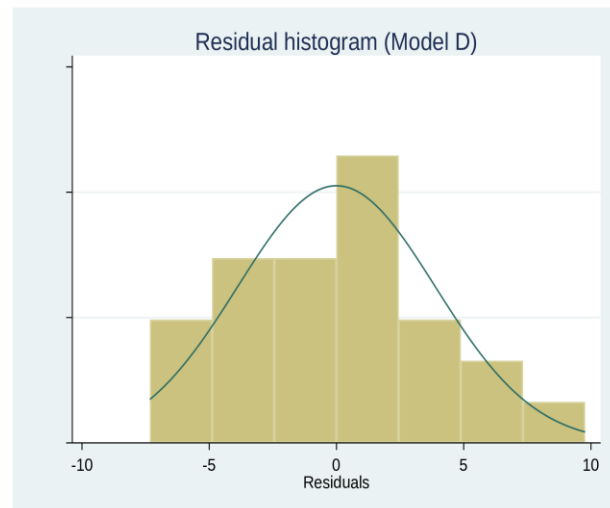
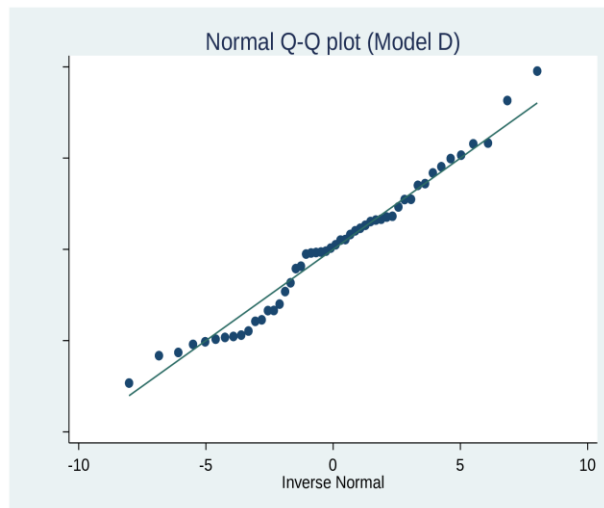


Figure 4. Normal Q–Q plot (Model D).

Figure 5. Residual histogram (Model D).

To evaluate influential observations, leverage values, studentized residuals, and Cook's distance are computed using an auxiliary OLS version of Model D. This OLS specification is used solely for diagnostic purposes, as influence measures are not directly defined within the GLM framework.

Following standard reference thresholds, observations are flagged if Cook's distance exceeds $4/N$ (0.078) or leverage exceeds $2p/N$ (0.392), where $N = 51$ and $p = 10$ (including the intercept and covariates). A small number of jurisdictions exhibit moderately elevated influence. Hawaii has the largest Cook's distance and a studentized residual exceeding 2 in absolute value, indicating that it is both an outcome outlier and moderately influential in model fit. Alaska, Delaware, and New Mexico also exceed the Cook's distance threshold, although with smaller magnitudes. Utah exhibits high leverage but relatively low Cook's distance, suggesting it is unusual in covariate space but does not meaningfully influence coefficient estimates. Importantly, all Cook's distance values remain well below conventional high-influence benchmarks (e.g., values near or above 1), indicating that no single jurisdiction dominates the results.

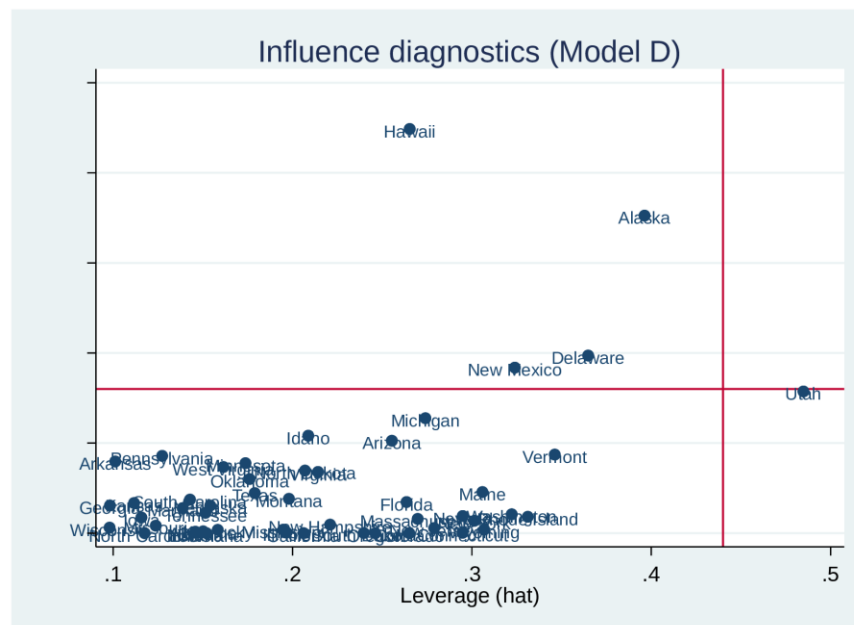


Figure 6. Influence diagnostics for the baseline OLS specification (Model D).

Note. The figure plots Cook's distance against leverage for each jurisdiction. Reference lines indicate common diagnostic thresholds for Cook's distance and leverage. States above these reference lines may be moderately influential, but no single observation appears to dominate the regression results.

In sum, the descriptive and diagnostic evidence indicates substantial cross-state variation in RN employment recovery and in structural state characteristics, including economic conditions, demographic composition, and policy environments. Raw comparisons suggest that Group 1 (Access & Affordability) is associated with higher average recovery, while Group 2 (Mandates) is associated with weaker recovery, and Group 3 shows no clear monotonic pattern. However, these patterns are descriptive and potentially confounded by cross-state differences in underlying conditions.

Correlation analysis shows weak bivariate associations between the outcome and policy indices, and VIF values indicate no serious multicollinearity concerns. Residual diagnostics suggest generally adequate model fit with mild tail deviations, and influence diagnostics identify a small number of moderately influential observations without evidence of a dominant jurisdiction.

Taken together, these results motivate the multivariable framework. The next section reports GLM estimates (Gaussian family, identity link) for Models A-D, along with OLS robustness checks, where RN employment recovery is regressed on the three policy-group indices and a set of state-level controls using heteroskedasticity-robust standard errors.

4.4 Regression Analysis

4.4.1 GLM Regression Results

Table 5. GLM Regression Results

VARIABLES	(1) Model A	(2) Model B	(3) Model C	(4) Model D
Group 1: Access & Affordability	1.113 (0.733)			0.914 (0.737)
Group 2: Public Health Mandates		-2.162* (1.263)		-2.025 (1.408)
Group 3: Labor/Care			0.067 (0.539)	0.130 (0.604)
Executives-to-RN pay ratio	-0.336 (1.427)	-0.522 (1.491)	-0.936 (1.484)	-0.189 (1.474)
Unemployment rate	0.229 (0.768)	1.643 (1.015)	0.373 (0.870)	1.326 (1.030)
GDP per capita (\$1,000s)	0.033 (0.061)	0.118* (0.069)	0.050 (0.069)	0.093 (0.070)
Housing price index	-0.001 (0.006)	0.002 (0.006)	0.000 (0.006)	0.001 (0.006)
Tax burden	-37.608* (21.278)	-9.058 (22.637)	-33.066 (24.286)	-16.847 (23.010)
Population age 65+ (%)	-0.432 (26.427)	8.939 (26.042)	-2.966 (27.052)	9.622 (25.702)
Constant	-0.544 (10.094)	-12.363 (10.831)	0.710 (10.915)	-11.115 (11.140)
Observations	51	51	51	51

Note. Entries are coefficients from generalized linear models with a Gaussian family and identity link. The dependent variable is percentage change in RN employment. Huber–White robust standard errors are reported in parentheses. Model A includes Group 1; Model B includes Group 2; Model C includes Group 3; Model D includes all three policy-group indices simultaneously. Inflation is omitted because it is constant across jurisdictions in the cross-sectional sample. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5 presents generalized linear model (GLM) estimates of RN employment recovery using a Gaussian family with an identity link. The dependent variable is the percentage change in RN employment across 51 jurisdictions. Models A–C estimate each policy group separately, while Model D includes all three policy-group indices simultaneously with state-level economic and demographic controls. All models report Huber–White robust standard errors.

Across specifications, the most consistent pattern is observed for Group 2 (Public Health Mandates). In the bivariate specification (Model B), the Group 2 index is negatively associated with RN employment recovery ($\beta = -2.162$, $SE = 1.263$, $p = 0.087$). In the fully specified model (Model D), the magnitude remains similar ($\beta = -2.025$, $SE = 1.408$), although the estimate becomes less precise and is not statistically significant at conventional levels. This suggests a stable negative directional relationship, but limited statistical precision in the cross-sectional setting.

The Group 1 (Access & Affordability) index is positively associated with RN employment recovery across specifications, though estimates are imprecise. In Model A, a one-unit increase in the index corresponds to a 1.11 percentage point increase in RN employment growth ($\beta = 1.113$, $SE = 0.733$). The effect remains positive in the full model ($\beta = 0.914$, $SE = 0.737$), indicating robustness in sign but not statistical significance. In contrast, the Group 3 (Labor Mobility & Care Disruption) index shows no meaningful association with RN employment recovery in any specification (Model C: $\beta = 0.067$, $SE = 0.539$; Model D: $\beta = 0.130$, $SE = 0.604$). The estimates are close to zero and statistically insignificant, suggesting no detectable net effect at the state level in this aggregate specification.

Overall, the GLM results indicate (i) a negative association for mandate-intensive policy environments, (ii) a positive but statistically imprecise association for access and affordability policies, and (iii) no meaningful association for labor mobility and care disruption policies. Importantly, none of the policy-group coefficients in the full model reach conventional significance levels; therefore, the results should be interpreted as directional evidence rather than causal estimates.

Control variables do not show strong or stable explanatory power across specifications. The executives-to-RN pay ratio is not statistically significant in any model. Unemployment and GDP per capita are positive but imprecisely estimated, suggesting broad macroeconomic differences across states rather than systematic structural effects. The Tax burden variable is negative in simpler specifications and approaches marginal significance in Model A, but this relationship weakens in the full model. Housing prices and the share of population aged 65+ remain small and statistically insignificant across all models.

In sum, once policy indices are included, state-level controls do not exhibit robust predictive power for RN employment recovery in this cross-sectional framework. The most consistent empirical patterns are the negative association for Group 2 and weak, imprecise associations for Groups 1 and 3, but these results should be interpreted cautiously given the limited sample size and cross-state heterogeneity.

4.4.2 GLM Predicted Values

Figures 7–9 plot adjusted predictions from the full GLM specification (Model D), holding all other covariates at their sample means. The three figures correspond to Group 1 (Access & Affordability), Group 2 (Public Health Mandates), and Group 3 (Labor Mobility & Care Disruption). Presenting adjusted predictions in this way complements the coefficient tables by showing the implied level differences in a more intuitive form.

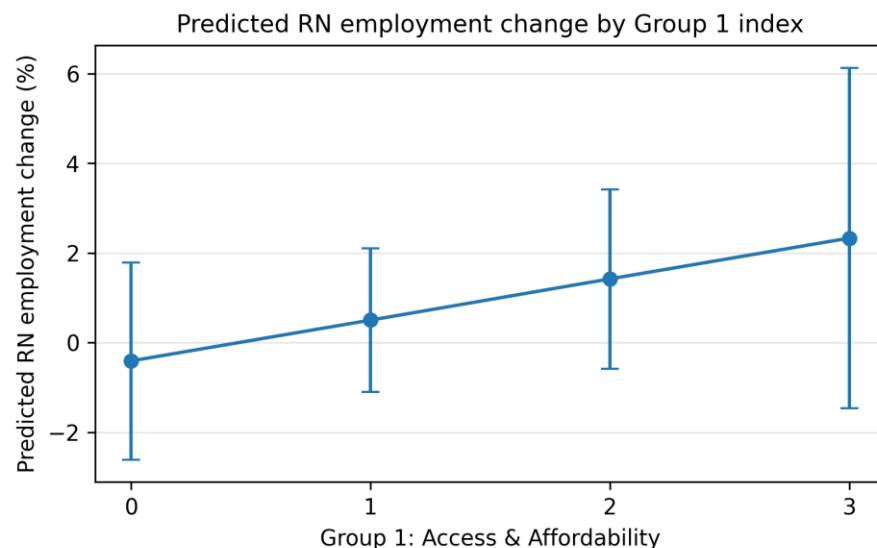


Figure 7. Predicted RN employment change by Group 1 (Access & Affordability) index.

Note. Predictions are based on Model D (Table 5), holding all covariates at sample means. Vertical bars represent 95% confidence intervals.

Figure 7 shows a generally increasing pattern in predicted RN employment recovery as the Group 1 index rises. The adjusted predictions move from negative values at low index levels to positive values at higher levels, indicating a modest upward gradient. This pattern is consistent with the positive coefficient estimates for Group 1 in Table 5. However, the confidence intervals are wide and overlap zero across much of the distribution, suggesting that the estimated differences are not statistically precise.

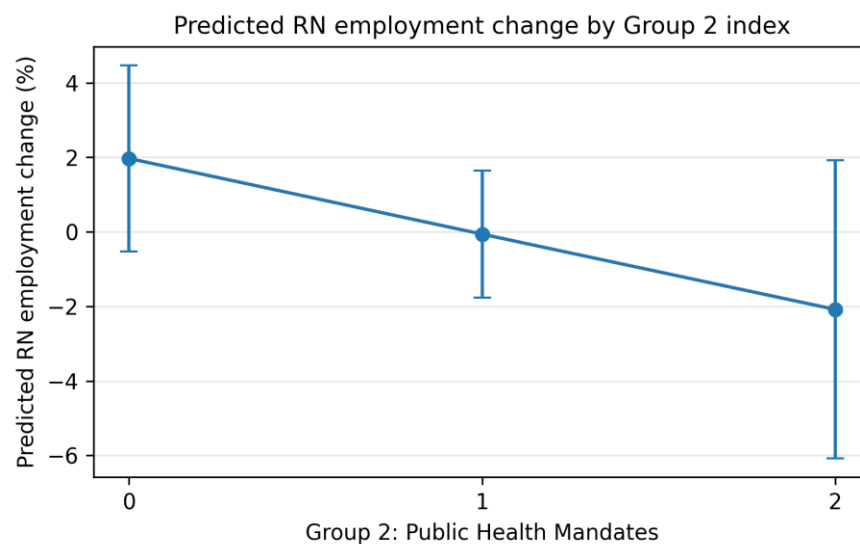


Figure 8. Predicted RN employment change by Group 2 (Public Health Mandates) index.

Note. Predictions are based on Model D (Table 5), holding all covariates at sample means. Vertical bars represent 95% confidence intervals.

Figure 8 displays a downward pattern in predicted RN employment recovery as the Group 2 index increases. Higher levels of mandate intensity are associated with lower predicted RN employment change, consistent with the negative coefficient on Group 2 in Table 5. As with Group 1, however, the confidence intervals are relatively wide, indicating substantial uncertainty around the predicted differences.

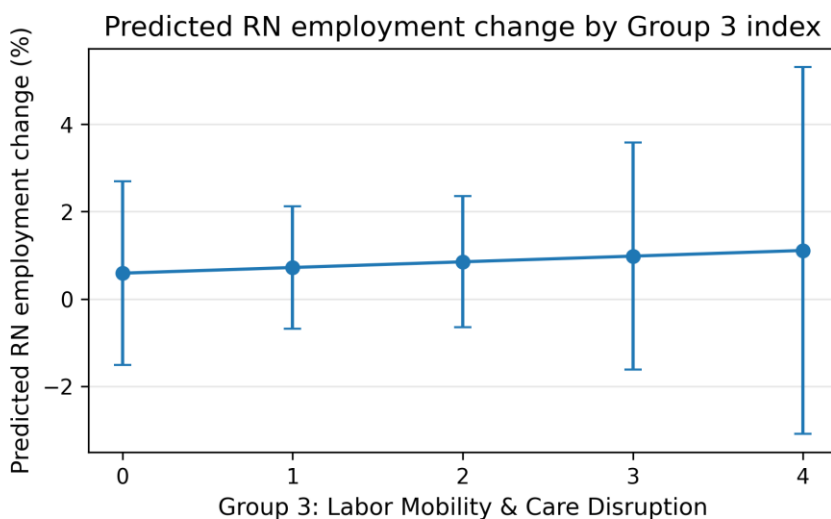


Figure 9. Predicted RN employment change by Group 3 (Labor Mobility & Care Disruption) index. Note. Predictions are based on Model D (Table 5), holding all covariates at sample means. Vertical bars represent 95% confidence intervals.

Figure 9 shows limited variation in predicted RN employment recovery across the observed range of the Group 3 index. The predicted values change only slightly as the index increases, and the confidence intervals overlap substantially at all levels. This is consistent with the near-zero coefficient estimates for Group 3 reported in Table 5, indicating little detectable net association in the aggregate specification.

Taken together, Figures 7–9 reinforce the regression results from Table 5. Group 1 exhibits a weak positive gradient, Group 2 shows a weak negative gradient, and Group 3 displays no meaningful variation in predicted outcomes. However, the wide confidence intervals across all three figures suggest that these patterns should be interpreted as directional relationships rather than precise estimates of policy effects.

4.4.3 Robustness Checks

To assess the robustness of the baseline findings, we conduct a series of specification checks examining sensitivity to alternative estimators and control variable definitions, as well as potential confounding from pre-existing differences in employment scale. The baseline model is estimated using a GLM with a Gaussian family and identity link (Table 5), while the primary robustness specification is an OLS model with Huber–White robust standard errors. Additional GLM specifications are used as sensitivity checks, including (i) replacement of the Tax burden measure with the 2021 tax variable and (ii) inclusion of lagged total employment.

Table 6 reports results across these specifications. The OLS baseline results serve as the primary robustness benchmark and confirm the baseline GLM findings. Across all models, the estimated coefficients on the three policy-group indices remain stable in sign, magnitude, and interpretation.

Table 6. Main Results and Robustness Checks (Model D Only)

Variable	OLS Base	GLM Tax (2021)	GLM Lagged Emp	OLS Lagged Emp
Group 1: Access & Affordability	0.914 (0.814)	0.926 (0.743)	0.875 (0.720)	0.875 (0.805)
Group 2: Public Health Mandates	-2.025 (1.555)	-1.997 (1.324)	-2.028 (1.403)	-2.028 (1.569)
Group 3: Labor/Care	0.130 (0.667)	0.282 (0.669)	0.118 (0.607)	0.118 (0.678)
Executives-to-RN pay ratio	-0.189 (1.628)	-0.175 (1.513)	-0.050 (1.439)	-0.050 (1.609)
Unemployment rate	1.326 (1.137)	1.359 (1.015)	1.418 (1.019)	1.418 (1.139)
GDP per capita (\$1,000s)	0.093 (0.077)	0.108 (0.068)	0.092 (0.071)	0.092 (0.079)
Housing price index	0.001 (0.007)	0.000 (0.007)	0.001 (0.006)	0.001 (0.007)
Tax burden	-16.847 (25.410)	-0.445 (0.463)	-15.531 (23.725)	-15.531 (26.525)
Population age 65+ (%)	9.622 (28.383)	16.409 (26.325)	7.922 (26.458)	7.922 (29.581)
Log lagged employment	—	—	-0.221 (0.615)	-0.221 (0.687)
Constant	-11.115 (12.302)	-10.181 (11.744)	-8.962 (13.687)	-8.962 (15.303)
Observations	51	50	51	51
R-squared	0.149	—	—	0.151

Note: Coefficients reported with standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Specifically, Group 1 (Access & Affordability) remains positive across all specifications, ranging from 0.875 to 0.926. Group 2 (Public Health Mandates) remains consistently negative, ranging from approximately -1.997 to -2.028. Group 3 (Labor Mobility and Care Disruption) remains close to zero and statistically weak across all specifications.

The GLM sensitivity checks show that substituting the alternative 2021 tax-burden measure produces nearly identical estimates relative to the baseline specification, indicating that the results are not driven by the choice of fiscal control variable. Similarly, the inclusion of lagged employment produces minimal changes in the estimated coefficients, and the lagged employment variable itself is negative but statistically insignificant, suggesting that pre-existing employment scale does not materially affect short-run RN employment growth.

Across all specifications, model fit remains stable, with R-squared values ranging from 0.149 to 0.151. Overall, the robustness results confirm that the main findings are not sensitive to estimator choice (OLS vs GLM), alternative control variable definitions, or baseline employment scale adjustments. However, given the relatively small cross-sectional sample of 51 jurisdictions, the results should be interpreted as directional rather than strictly causal.

5. DISCUSSION OF MAIN RESULTS

This study examines whether bundles of state-level COVID-19 policies adopted in 2021 are associated with registered nurse (RN) employment recovery from 2021 to 2022. Rather than estimating isolated policy effects, the analysis constructs three composite indices—Access & Affordability (Group 1), Public Health Mandates (Group 2), and Labor Mobility & Care Disruption (Group 3)—to reflect the bundled nature of state

policy responses during the pandemic. This approach provides a parsimonious framework for analyzing multi-dimensional policy environments in a small cross-section of U.S. states.

Across the baseline GLM specification and a series of robustness checks, three consistent patterns emerge. First, the Access & Affordability index is positively associated with RN employment recovery across all specifications. The magnitude is economically meaningful, suggesting that additional access-oriented policies are associated with roughly one percentage point higher RN employment growth, although estimates are imprecisely estimated with wide confidence intervals. This pattern is consistent with a continuity-of-care interpretation, whereby policies such as telehealth expansion and reduced cost barriers may support service utilization and stabilize healthcare labor demand during the recovery period.

Second, the Public Health Mandates index is consistently negatively associated with RN employment recovery, with effect sizes centered around approximately -2 percentage points per additional mandate. However, this relationship should not be interpreted as a direct causal effect of mandates on employment outcomes. A more plausible interpretation is policy endogeneity: states experiencing more severe pandemic conditions may have simultaneously adopted more stringent public health mandates and experienced weaker subsequent employment recovery. Under this interpretation, the coefficient reflects underlying crisis severity and labor-market stress rather than a direct effect of policy restrictiveness.

Third, the Labor Mobility & Care Disruption index exhibits no systematic association with RN employment recovery. Coefficients are small in magnitude and statistically indistinguishable from zero across all specifications. This null result is consistent with offsetting mechanisms within the index: while policies such as paid sick leave may improve labor-force attachment, school closures and caregiving disruptions may reduce available nursing labor supply. The net effect may therefore be close to zero in aggregate employment growth, even if important effects exist along other labor-market margins such as hours worked, turnover, or burnout.

Overall, the results indicate that policy environment structure is associated with variation in RN employment recovery, but the estimated relationships are modest in precision and should be interpreted as directional rather than causal. The findings also highlight the limitations of short-horizon cross-sectional employment growth as an outcome measure for capturing structural labor-market adjustment in healthcare systems.

5.1 Managerial Implications

Although the estimates are not causal, the directional patterns offer several practical implications for hospital leaders, health-system managers, and state workforce planners. First, the positive sign on Access & Affordability suggests that continuity-of-care mechanisms, such as telehealth capacity, medication access, and reduced patient cost barriers, may help preserve service utilization and stabilize staffing needs during major disruptions. Managers can therefore treat digital access, flexible care channels, and patient affordability as part of workforce-resilience planning rather than as separate administrative functions. Second, the negative association for mandate-intensive environments should not be interpreted as evidence that mandates reduce RN employment. Because mandate adoption may reflect underlying outbreak severity and operational stress, managers can instead view a high-intensity public-health environment as a signal to activate contingency staffing, cross-training, retention efforts, and surge-capacity planning. Third, the near-zero net relationship for the Labor Mobility & Care Disruption index suggests that aggregate headcount can conceal offsetting workforce pressures. Health systems should therefore monitor absenteeism, hours worked, turnover, vacancies, and caregiving constraints in addition to total RN employment. Overall, the findings support flexible, locally adapted workforce strategies rather than reliance on any single policy bundle.

5.2 Limitations and Contemporary Relevance

This study has several limitations. First, the cross-sectional sample contains only 51 jurisdictions, limiting statistical power and contributing to wide confidence intervals. Second, policy adoption is likely endogenous to pandemic severity, healthcare-system stress, political conditions, and pre-existing workforce characteristics; accordingly, the estimates are associations rather than causal effects. Third, the analysis cannot fully control for pre-pandemic nursing workforce conditions such as RN density, vacancy rates,

hospital capacity, or local staffing shortages. The lagged-employment specification partially adjusts for scale but does not capture all baseline differences in workforce tightness. Fourth, the additive policy indices simplify heterogeneous policies and do not capture differences in timing, intensity, enforcement, compliance, or the interaction among individual policies. Finally, the outcome measures net RN employment change over a short horizon and may obscure turnover, vacancies, labor-force exit, migration, changes in hours worked, and other adjustment margins that matter to healthcare organizations.

The findings should also be interpreted in light of the substantially different post-pandemic operating environment. The federal COVID-19 public health emergency ended on May 11, 2023, and many emergency flexibilities expired or were replaced by post-emergency rules (U.S. Department of Health and Human Services, Office of Inspector General, 2023). Telehealth remains part of healthcare delivery, but Medicare telehealth services are now managed through an evolving post-emergency policy framework rather than the same emergency arrangements that applied in 2021 (Centers for Medicare & Medicaid Services, 2026). At the same time, the RN labor market continues to face strong structural demand. The U.S. Bureau of Labor Statistics reports about 3.39 million RN jobs in 2024 and projects 5% employment growth from 2024 to 2034, with about 189,100 openings per year on average, many reflecting replacement needs as workers leave the occupation or labor force (U.S. Bureau of Labor Statistics, 2025). These changes mean that the associations estimated for 2021-2022 should not be interpreted as direct predictions for current RN employment. Current workforce outcomes are increasingly shaped by longer-run replacement needs, an aging population, chronic-care demand, compensation and staffing practices, and the normalization of virtual care. The study is therefore best viewed as evidence about RN workforce recovery under an exceptional policy environment, with implications for future shocks and emergency preparedness rather than as a model of the current labor market.

6. CONCLUSION

This study provides descriptive evidence on the relationship between state-level COVID-19 policy bundles and RN employment recovery during 2021-2022. Across the baseline GLM and robustness specifications, Access & Affordability policies show a consistently positive but imprecise association with RN employment growth, Public Health Mandates show a consistently negative but imprecise association, and Labor Mobility & Care Disruption shows little net relationship. The results are directional rather than causal, but they demonstrate the value of considering policy environments as bundles when examining healthcare workforce recovery.

The practical contribution is similarly cautious. For healthcare managers and workforce planners, the results emphasize continuity-of-care capacity, flexible workforce planning, and monitoring of staffing margins beyond total headcount. Because the policy and labor-market environment has changed substantially since 2021, the findings are most useful for understanding workforce adjustment during a major disruption and for informing preparedness for future shocks.

Future research could extend the analysis using multi-year panel data and stronger identification strategies, such as difference-in-differences or event-study designs, and could incorporate richer workforce outcomes including vacancy rates, wage adjustments, turnover, hours worked, staffing ratios, and interstate migration. Such extensions would provide a more complete picture of how healthcare labor markets respond to policy changes over time.

REFERENCES:

- Adolph, C., Amano, K., Bang-Jensen, B., Fullman, N., Magistro, B., Reinke, G., et al. (2022). The pandemic policy U-turn: Partisanship, public health, and race in decisions to ease COVID-19 social distancing policies in the United States. *Perspectives on Politics*, 20(2), 595–617.
- Alexander, G. C., & Qato, D. M. (2020). Ensuring access to medications in the US during the COVID-19 pandemic. *JAMA*. Advance online publication. <https://doi.org/10.1001/jama.2020.6016>

- Amuedo-Dorantes, C., Arenas-Arroyo, E., & Sevilla, A. (2023). COVID-19 and the gender gap in the labor market. *ILR Review*, 76(1), 56–85.
- Bloom, M. (1999). The performance effects of pay dispersion on individuals and organizations. *Academy of Management Journal*, 42(1), 25–40.
- Buchan, J., & Aiken, L. (2008). Solving nursing shortages: A common priority. *Journal of Clinical Nursing*, 17(24), 3262–3268.
- Buerhaus, P. I., Staiger, D. O., Auerbach, D. I., Yates, M. C., & Donelan, K. (2022). Nurse employment during the first fifteen months of the COVID-19 pandemic. *Health Affairs*, 41(1), 79–85.
- Capano, G., & Howlett, M. (2021). Causal logics and mechanisms in policy design. *Public Policy and Administration*, 36(2), 141–162.
- Centers for Medicare & Medicaid Services. (2026). Telehealth. U.S. Department of Health and Human Services. <https://www.cms.gov/medicare/coverage/telehealth>
- Cebula, R. J., & Clark, J. R. (2011). Migration, economic freedom, and personal freedom. *The Journal of Private Enterprise*, 27(1), 43–62. <http://www.econis.eu/PPNSET?PPN=726581097>.
- Cebula, R., Foley, M., Shaw, J., Chin, K., & Ma, G. (2025). Telehealth and healthcare workforce recovery in the U.S.: State-level evidence on labor market stabilization during COVID-19 (Working Paper).
- Chetty, R., Friedman, J. N., Hendren, N., & Stepner, M. (2020). How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data (NBER Working Paper No. 27431).
- Chua, K.-P., Conti, R. M., & Becker, N. V. (2022). Trends in and factors associated with out-of-pocket spending for COVID-19 hospitalizations in 2020 and 2021. *JAMA Network Open*, 5(2), e2148237.
- Coibion, O., Gorodnichenko, Y., & Weber, M. (2020). Labor markets during the COVID-19 crisis: A preliminary view (NBER Working Paper No. 27017). National Bureau of Economic Research.
- Collins, C., Landivar, L. C., Ruppanner, L., & Scarborough, W. J. (2021). COVID-19 and the gender gap in work hours. *Gender, Work & Organization*, 28(S1), 101–112.
- Cowherd, D. M., & Levine, D. I. (1992). Product quality and pay equity between lower-level employees and top management: An investigation of distributive justice theory. *Administrative Science Quarterly*, 37(2), 302–320.
- Demeke, H. B., Merali, S., Marks, S., Pao, L. Z., Romero, L., Sandhu, P., et al. (2021). Trends in use of telehealth among health centers during the COVID-19 pandemic—United States, June 26–November 6, 2020. *MMWR Morbidity and Mortality Weekly Report*, 70, 240–244.
- Dolan, R. (2020, April 30). States are shifting how they cover prescription drugs in response to COVID-19. KFF.
- Foley, M., Shaw, J., Sockbeson, C., Chin, K., & Ma, G. (2026). State COVID-19 Policy Configurations and Registered Nurse Employment Recovery: A Mixed-Set Qualitative Comparative Analysis of U.S. States and the District of Columbia (2021–2022). *American Journal of Economics and Sociology*, 85-3, 389-403, DOI: 10.1111/ajes.70043
- Ganong, P., & Shoag, D. (2017). Why has regional income convergence in the U.S. declined? *Journal of Urban Economics*, 102, 76–90.

- Garcia, K. S. D., & Cowan, B. W. (2022). The impact of school and childcare closures on labor market outcomes during the COVID-19 pandemic (NBER Working Paper No. 29641).
- Goldin, C. (2022). Understanding the economic impact of COVID-19 on women. Brookings Papers on Economic Activity, Spring 2022.
- Goolsbee, A., & Syverson, C. (2021). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics*, 193, 104311.
- Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., et al. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5, 529–538.
- Hollander, J. E., & Carr, B. G. (2020). Virtually perfect? Telemedicine for COVID-19. *New England Journal of Medicine*, 382(18), 1679–1681.
- Karaca-Mandic, P., Georgiou, A., & Sen, A. (2021). Healthcare utilization and spending during the COVID-19 pandemic. *Health Affairs*, 40(3), 445–452.
- Landivar, L. C., Scarborough, W. J., Ruppanner, L., Collins, C., & Rouse, L. (2023). Remote schooling and mothers' employment during the COVID-19 pandemic. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 9(3), 134–158.
- Lee, E. S. (1966). A theory of migration. *Demography*, 3(1), 47–57.
- Patel, S. Y., Rose, S., Barnett, M. L., Huskamp, H. A., Uscher-Pines, L., & Mehrotra, A. (2021). Community factors associated with telemedicine use during the COVID-19 pandemic. *JAMA Network Open*, 4(5), e2110330.
- Pichler, S., Wen, K., & Ziebarth, N. R. (2020). COVID-19 emergency sick leave has helped to flatten the curve in the United States. *Health Affairs*, 39(12), 2197–2204.
- Shaw, J. D., Gupta, N., & Delery, J. E. (2002). Pay dispersion and workforce performance: Moderating effects of incentives and interdependence. *Strategic Management Journal*, 23(6), 491–512.
- U.S. Bureau of Labor Statistics. (2025). Registered nurses. *Occupational Outlook Handbook*. U.S. Department of Labor. <https://www.bls.gov/ooh/healthcare/registered-nurses.htm>
- U.S. Department of Health and Human Services, Office of Inspector General. (2023). *OIG's COVID-19 public health emergency flexibilities end on May 11, 2023 upon expiration of the COVID-19 public health emergency declaration*. <https://oig.hhs.gov/reports/featured/coronavirus/covid-flex-expiration/>
- Wang, Y., et al. (2025). Impacts of state COVID-19 vaccine mandates for health care workers on health sector employment. *American Journal of Public Health*, 115(2), 214–222.
- Wosik, J., Fudim, M., Cameron, B., Gellad, Z. F., Cho, A., Phinney, D., et al. (2020). Telehealth transformation: COVID-19 and the rise of virtual care. *Journal of the American Medical Informatics Association*, 27(6), 957–962.
- World Health Organization. (2020). *State of the world's nursing 2020: Investing in education, jobs and leadership*. Geneva: WHO.

HOW BANK SOCIAL MEDIA ACTIVITIES SHAPE BRAND EQUITY: EVIDENCE FROM CUSTOMERS OF COMMERCIAL BANKS IN GHANA

Fidelis Quansah, University of Professional Studies Accra, Ghana
Timothy Kwabla Zilevu, University of Professional Studies Accra, Ghana
Nathan Kobina Austin, University of the Virgin Islands, U.S.A.
Julia Tomety, University of Professional Studies Accra, Ghana

[dx.doi.org/10.18374/JABE-26-3.4](https://doi.org/10.18374/JABE-26-3.4)

ABSTRACT

This paper examines how bank social media marketing activities (SMMAs) shape customer-based brand equity (CBBE). Using a quantitative cross-sectional design, structured questionnaires measuring five SMMA dimensions, brand experience, brand trust, and four CBBE dimensions were administered to 600 customers at the banking halls of all 23 licensed commercial banks in Ghana, yielding 534 (89.0%) responses following systematic data cleaning. In stage 1 of the study, eWOM ($\beta = 0.709$) and trendiness ($\beta = 0.501$) were found to be the strongest predictors of brand experience, while customization ($\beta = 0.418$) was the dominant predictor of brand trust. In Stage 2, brand trust was a stronger predictor of brand awareness ($\beta = 0.530$) and perceived quality ($\beta = 0.517$) than brand experience, while brand experience was the stronger predictor of brand loyalty ($\beta = 0.548$) and both brand experience and brand trust mediated 16 of 20 tested paths.

Keywords: social media marketing, brand equity, brand experience, and brand trust.

1. INTRODUCTION

The rapid evolution of digital technologies has fundamentally reshaped how financial institutions interact with their customers, with social media emerging as one of the most influential marketing communication channels in the 21st century (Santini et al., 2020). Commercial banks, which operate in a highly competitive and service-driven environment, have increasingly integrated social media platforms into their branding and customer engagement strategies (Vives, 2019). Understanding how these activities contribute to customer-based brand equity (CBBE) formation is of both academic and practical significance, particularly in emerging economies where digital adoption continues to accelerate (Bruce et al., 2023).

Brand equity, defined as the differential value a brand confers upon a product or service, is widely recognized as a critical intangible asset for banks (Keller, 1993). In the banking sector, where core service offerings are largely undifferentiated, strong brand equity serves as a key source of competitive advantage, influencing customer acquisition, retention, and long-term profitability (Nguyen, 2021). Customer-based brand equity, as conceptualized by Keller (1993) and Aaker (1991), encompasses brand awareness, brand associations, perceived quality, and brand loyalty, all of which can be shaped by a firm's marketing communications (Tasci, 2020).

Social media marketing activities (SMMAs) have attracted growing scholarly attention as determinants of brand equity across industries (Godey et al., 2016). Hafez (2021) found that SMMAs did not directly influence brand equity in Bangladeshi banking but operated through brand love as a full mediator and brand trust as a partial mediator. Garanti and Kissi (2019) demonstrated in Latvia that social media brand personality traits contributed significantly to brand equity. These studies collectively establish that the social media-brand equity relationship in banking is driven by affective and relational mechanisms rather than direct exposure, but they also reveal that the specific mediating mechanisms differ substantially across contexts.

Despite this growing body of evidence, three gaps persist. First, empirical work on this relationship in banking is concentrated in South Asian, Eastern European, and Southern African contexts, with no study examining it specifically in West Africa or Ghana. Second, the pathways through which distinct SMMA dimensions individually shape CBBE in a West African banking context remain untested. Third, the sequential mediation model positioning brand experience and brand trust as organism-level states within an S-O-R framework has not been applied to bank brand equity formation in sub-Saharan Africa. Ghana presents a compelling setting: its 2017-2019 banking crisis, which resulted in the revocation of nine bank licenses by the Bank of Ghana, fundamentally altered the institutional trust landscape, creating conditions under which the mechanisms through which social media marketing builds brand equity may differ from those observed in more stable banking environments.

This study argues that in post-crisis banking markets, social media brand equity formation depends less on entertainment reach and more on trust-building signals such as content customization, peer validation, and institutional credibility. Testing this argument, the study examines how SMMA dimensions shape CBBE among customers of commercial banks in Ghana through the sequential mediation of brand experience and brand trust. The findings contribute to S-O-R-based brand equity theory by demonstrating how post-crisis institutional context conditions the relative importance of experiential versus trust-based mediating mechanisms.

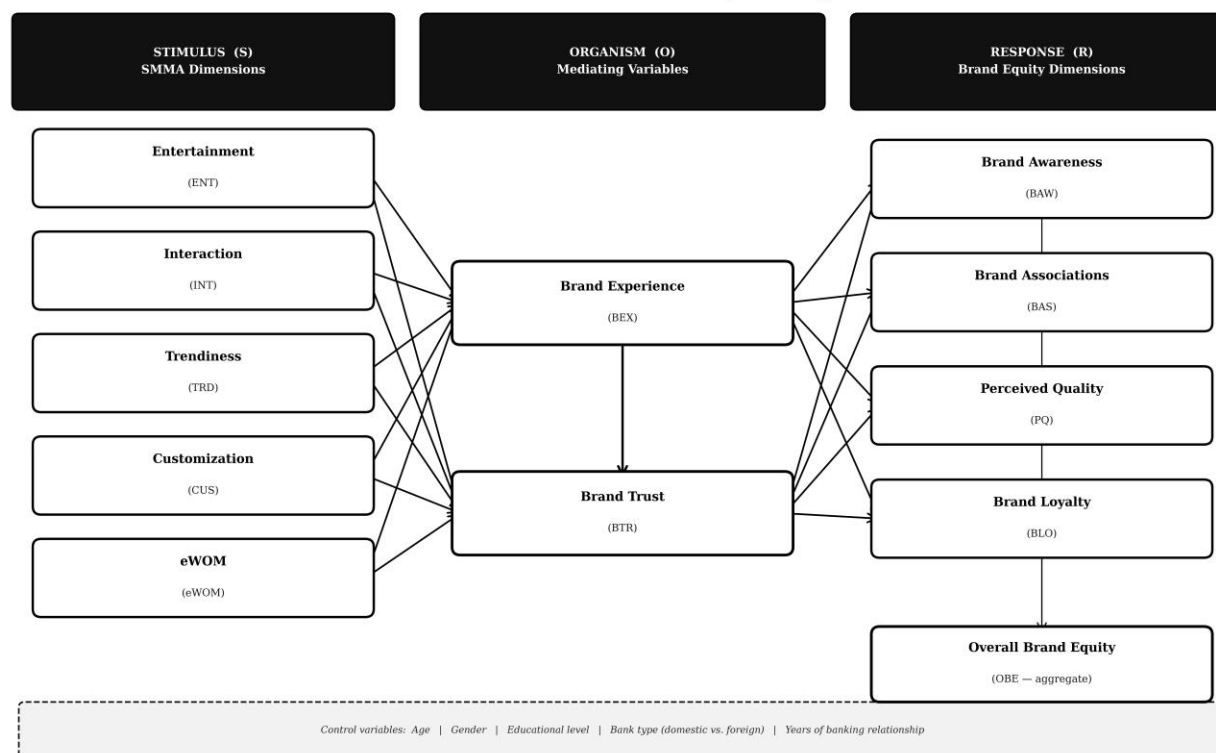
2. THEORETICAL FRAMEWORK AND HYPOTHESES DEVELOPMENT

2.1 *The S-O-R Framework and the Study Model*

This study is anchored in the Stimulus-Organism-Response (S-O-R) framework developed by Mehrabian and Russell (1974), which proposes that external stimuli act upon an organism's internal states, which in turn produce observable behavioral and evaluative responses (Hochreiter et al., 2023). The framework has been extensively applied in social media marketing research to explain how brand-related stimuli transmitted online influence consumer perceptions and brand outcomes (Koay et al., 2020).

The study model adopts a two-stage architecture. In Stage 1, the five SMMA dimensions act as stimuli that influence brand experience and brand trust as organism-level states. Stage 1 also incorporates the sequential path from brand experience to brand trust, consistent with evidence that brand experience is an antecedent of trust-based brand evaluations (Moedeen et al., 2023). In Stage 2, brand experience and brand trust jointly predict the four CBBE dimensions as response outcomes. This two-stage structure cleanly separates the stimulus-to-organism pathways from the organism-to-response pathways, avoiding the model overspecification that arises when direct IV-to-DV paths are included alongside multiple correlated mediators. Figure 1 presents the conceptual model.

Figure 1. Conceptual Model: How Bank Social Media Activities Shape Brand Equity
Theoretical S-O-R Framework — Prior to Empirical Testing



Source: authors own work

Figure 1. Conceptual Model: Bank Social Media Activities and Brand Equity (S-O-R Framework — Prior to Testing)

Koay et al. (2020) applied this architecture to the SMMA-CBBE relationship and found that brand experience mediates the stimulus-to-response pathway. Malarvizhi et al. (2022) used the same framework to connect SMMA dimensions to brand awareness and brand loyalty. Chavadi et al. (2023) demonstrated that social media brand community stimuli produce significant effects on all four CBBE dimensions through organism-level processing in a banking-specific context.

2.2 SMMA Dimensions and Brand Experience (H1a-H1e)

Brand experience refers to the sensory, affective, behavioral, and intellectual responses evoked by brand-related stimuli (Brakus et al., 2009). Koay et al. (2020) confirmed that perceived SMMA produce significant effects on brand experience. Hafez (2022a) found that brand experience fully mediates the SMMA-to-brand equity relationship in Bangladeshi banking. Dananjoyo et al. (2025) identified interactivity, responsiveness, and information quality as key SMMA determinants of brand experience. Therefore **H1a:** Entertainment positively influences brand experience, **H1b:** Interaction positively influences brand experience, **H1c:** Trendiness positively influences brand experience, **H1d:** Customization positively influences brand experience and **H1e:** eWOM positively influences brand experience.

2.3 SMMA Dimensions and Brand Trust (H2a-H2e)

Brand trust reflects customers' confidence in a brand's reliability, integrity, and benevolence (Morgan and Hunt, 1994). Althuaini (2022) found that customization, entertainment, and promotional social media activities had the strongest effects on brand trust among Saudi Arabian bank customers. Hafez (2021) positioned brand trust as a partial mediator in the SMMA-brand equity relationship in Bangladesh. Rico et al. (2024) confirmed trust as the critical intermediary between social media dimensions and brand loyalty in Colombian banking. Therefore **H2a:** Entertainment positively influences brand trust, **H2b:** Interaction positively influences brand trust, **H2c:** Trendiness positively influences brand trust, **H2d:** Customization positively influences brand trust and **H2e:** eWOM positively influences brand trust.

2.4 Brand Experience and Brand Trust (H3)

Brand experience produces emotional and cognitive responses that shape relational brand perceptions, including trust (Brakus et al., 2009). Hafez (2022a) confirmed that brand experience is an antecedent of brand trust in the SMMA-bank brand equity model. Moedeen et al. (2023) found that brand experience sequentially strengthens brand trust, which ultimately drives brand equity. Therefore **H3**: Brand experience positively influences brand trust.

2.5 Mediating Roles of Brand Experience and Brand Trust (H4, H5)

The SMMA-brand equity relationship in banking operates primarily through affective and relational mechanisms rather than through direct exposure (Hafez, 2021, 2022a). Brand experience mediates the pathway to brand associations and perceived quality (Koay et al., 2020; Dananjoyo et al., 2025). Brand trust mediates pathways to brand awareness, perceived quality, and brand loyalty (Althuwaini, 2022; Elareshi et al., 2023; Rico et al., 2024). Therefore **H4a-H4e**: Brand experience mediates the relationship between each SMMA dimension and each CBBE dimension and **H5a-H5e**: Brand trust mediates the relationship between each SMMA dimension and each CBBE dimension.

3. METHODOLOGY

3.1 Research Design and Context

A quantitative cross-sectional survey design was adopted within a positivist philosophical tradition, consistent with comparable banking-sector SMMA studies (Hafez, 2021; Althuwaini, 2022). The research context is Ghana's commercial banking sector. Following the deregulation of the banking system in 2006, which introduced a universal banking regime and intensified competitive pressures (Amoako et al., 2017), Ghana's banks subsequently experienced a systemic crisis. Between 2017 and 2019, the Bank of Ghana revoked the operating licenses of nine commercial banks due to capital adequacy failures and corporate governance deficiencies. This episode fundamentally altered the trust landscape in Ghanaian banking and provides the institutional backdrop that distinguishes the Ghanaian context from the South Asian and European settings in which most prior SMMA-brand equity research was conducted. Ghana's rising smartphone penetration and social media adoption rates (Mohammed and Donkoh, 2024) make it a timely setting for this investigation.

3.2 Sample and Data Collection

The target population was customers of all 23 licensed commercial banks operating in Ghana at the time of data collection, as listed in the Bank of Ghana's official register. A purposive criterion was applied: all respondents were required to hold an active account at the selected bank and to actively follow their bank's official social media account on at least one platform. This criterion was verified through a two-item screener at the start of each questionnaire. Data were collected in person by trained research assistants who administered structured questionnaires directly to customers at the banking halls and customer service areas of all 23 banks, covering branches in Greater Accra, Ashanti, Western, Eastern, and Central regions to ensure geographic coverage beyond Accra. Questionnaires were distributed proportionally based on each bank's estimated average monthly customer footfall, which was obtained from branch managers prior to fieldwork. Customer interception was systematic: every third customer entering the banking hall during operating hours was approached. Of 600 questionnaires distributed over the period of March 2025 till January 2026, questionnaires were cleaned for completeness, internal consistency, and straight-lining. Sixty-six questionnaires were removed, yielding 534 usable responses and an 89.0% usable response rate. A pilot test with 30 bank customers from two banks preceded the main fieldwork, leading to minor wording adjustments in two items. The sample of 534 exceeds the minimum of 200 cases for structural equation modeling (Hair et al., 2019) and satisfies the 10-observations-per-indicator rule for PLS-SEM (Kock and Hadaya, 2018).

3.3 Measures

All items were adapted from validated scales. The five SMMA dimensions were measured using four items each from Kim and Ko (2012). Brand experience was measured with nine items from Brakus et al. (2009) and Hafez (2022a), covering sensory, affective, behavioral, and intellectual dimensions. Brand trust was measured with eight items adapted from Morgan and Hunt (1994) and Althuwaini (2022). The four CBBE

dimensions were measured using items from Aaker (1991) and Yoo et al. (2000): brand awareness (three items), brand associations (two items), perceived quality (four items), and brand loyalty (five items). All items were rated on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). CBBE is reported as four reflective first-order dimensions rather than a single composite score, consistent with the Keller (1993) and Aaker (1991) conceptualization used in prior banking-sector SMMA studies (Hafez, 2021; Garanti and Kissi, 2019).

3.4 Analytical Approach

The measurement model was assessed for reliability (Cronbach's alpha, CR), convergent validity (AVE), and discriminant validity using HTMT ratios with bootstrapped 95% confidence intervals (5,000 resamples; Henseler et al., 2015). Outer loadings were inspected for all items to confirm item-level reliability above 0.70. The structural model was estimated in two stages consistent with the S-O-R architecture and the PLS-SEM tradition (Hair et al., 2019): Stage 1 estimated paths from the five SMMA dimensions to brand experience and brand trust (with brand experience included as a predictor of brand trust); Stage 2 estimated paths from brand experience and brand trust to the four CBBE dimensions. This two-stage approach avoids the model overspecification and suppression effects that arise from simultaneously including highly correlated SMMA predictors alongside their mediators in a single equation predicting CBBE outcomes. Path coefficients were derived from OLS regression on standardized variables (Rigdon, 2016). VIF was computed for each stage separately; all Stage 1 VIF values fall below 4.7 and all Stage 2 VIF values equal 1.26, confirming no multicollinearity concern at either stage. Effect sizes were computed as Cohen's f^2 and predictive relevance was assessed via Q^2 using ten-fold cross-validation. Mediation was assessed using standardized product-of-coefficients indirect effects aligned with the two-stage path model, with significance evaluated using 95% confidence intervals (Preacher and Hayes, 2008). Harman's single factor test confirmed that common method bias is not a substantial concern (first factor = 42.7% of variance, below the 50% threshold; Podsakoff et al., 2003). Non-response bias was assessed by comparing respondent demographic profiles across the 23 bank locations and verifying that respondents from different branches did not differ systematically on key demographic variables (Armstrong and Overton, 1977).

AI-assisted language editing disclosure: ChatGPT by OpenAI was used only for English language correction, grammar refinement, and clarity improvement. No AI tool was used to generate the research design, data, analysis, findings, interpretations, or conclusions. The authors reviewed all edits and remain fully responsible for the final manuscript.

For the submission checkbox, you can safely confirm:

The manuscript was created by the author(s), not by an AI tool or Large Language Model. AI was used only for English language correction and clarity refinement, and this use has been disclosed in the Methods and Acknowledgements.

4. RESULTS

4.1 Descriptive Statistics

Descriptive statistics across the 11 constructs confirm suitable data properties for SEM analysis.

Descriptive statistics across the 11 constructs ($N = 534$) show that brand awareness returned the highest mean ($M = 4.101$, $SD = 0.869$) and eWOM the lowest ($M = 2.888$, $SD = 1.038$). Skewness values fall between -1 and $+1$ and kurtosis values remain within acceptable ranges, confirming that the data do not severely violate parametric assumptions. ENT: $M = 3.285$, $SD = 0.824$; INT: $M = 3.412$, $SD = 1.024$; TRD: $M = 3.496$, $SD = 0.807$; CUS: $M = 3.066$, $SD = 0.800$; BEX: $M = 3.228$, $SD = 0.738$; BTR: $M = 3.611$, $SD = 0.911$; BAS: $M = 3.550$, $SD = 0.891$; PQ: $M = 3.549$, $SD = 0.858$; BLO: $M = 3.325$, $SD = 0.936$. All items were rated on a five-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree).

4.2 Measurement Model

Reliability and convergent validity indices are reported below. All outer loadings exceed 0.70, ranging from 0.780 to 0.942, confirming item-level reliability. Cronbach's alpha values range from 0.796 to 0.966, all exceeding 0.70 (Nunnally, 1978). CR values range from 0.902 to 0.966. AVE values range from 0.675 to

0.877, all exceeding 0.50, confirming convergent validity (Fornell and Larcker, 1981). Each construct's square root of AVE exceeds its highest inter-construct correlation, providing additional convergent validity support.

Construct-level reliability and convergent validity are strong throughout. Outer loadings for all indicators range from 0.780 to 0.942, all exceeding the 0.70 threshold. Cronbach's alpha ranges from 0.796 (BAS) to 0.966 (eWOM). Composite reliability (CR) ranges from 0.902 to 0.966. AVE ranges from 0.675 (BEX) to 0.877 (eWOM), all above 0.50 (Fornell and Larcker, 1981). By construct: ENT (loadings 0.835–0.894, α = 0.885, CR = 0.921, AVE = 0.744); INT (0.918–0.934, 0.947, 0.962, 0.862); TRD (0.826–0.894, 0.895, 0.927, 0.762); CUS (0.808–0.851, 0.854, 0.902, 0.698); eWOM (0.930–0.942, 0.953, 0.966, 0.877); BEX (0.780–0.860, 0.940, 0.949, 0.675); BTR (0.834–0.895, 0.949, 0.957, 0.736); BAW (0.865–0.894, 0.852, 0.911, 0.773); BAS (0.911–0.911, 0.796, 0.907, 0.831); PQ (0.877–0.902, 0.914, 0.940, 0.797); BLO (0.838–0.917, 0.938, 0.953, 0.802). Each construct's square root of AVE exceeds its highest inter-construct correlation, providing additional convergent validity support.

The inter-construct correlation matrix and the full HTMT matrix (Table 4) provide the basis for assessing discriminant validity. Of 55 construct pairs, 53 return HTMT ratios below the 0.85 threshold, confirming discriminant validity (Henseler et al., 2015). Two pairs marginally exceed the threshold: BAS-PQ (0.891) and PQ-BLO (0.878). Bootstrapped 95% confidence intervals (5,000 resamples) were computed for both pairs. Both intervals exclude 1.0: BAS-PQ [0.851, 0.931] and PQ-BLO [0.848, 0.907], confirming that these constructs, while theoretically proximate within the CBBE framework (Tasci, 2020), remain empirically distinguishable.

The inter-construct correlation matrix (lower triangle, Pearson r) shows that all bivariate correlations between the SMMA dimensions and the mediators are positive, ranging from $r = 0.065$ (CUS–BEX) to $r = 0.796$ (eWOM–BEX). The highest inter-predictor correlation is TRD–ENT ($r = 0.625$), and the highest mediator–outcome correlation is eWOM–BLO ($r = 0.758$). The full matrix confirms that no bivariate correlation is so high as to suggest construct redundancy among the SMMA dimensions prior to simultaneous regression.

The full HTMT matrix (55 construct pairs) confirms discriminant validity throughout the model. Of 55 pairs, 53 return HTMT ratios below 0.85 (Henseler et al., 2015). Selected values illustrating the dimensional structure: ENT–INT = 0.457, ENT–TRD = 0.703, ENT–CUS = 0.211, ENT–eWOM = 0.587, INT–TRD = 0.618, INT–CUS = 0.558, TRD–CUS = 0.439, BEX–BTR = 0.478, BEX–eWOM = 0.842, BTR–BAW = 0.640. The two pairs that marginally exceed 0.85 are BAS–PQ (HTMT = 0.891, 95% CI [0.851, 0.931]) and PQ–BLO (HTMT = 0.878, 95% CI [0.848, 0.907]). Both bootstrapped confidence intervals exclude 1.0, confirming that these theoretically proximate CBBE constructs remain empirically distinguishable (Tasci, 2020).

Harman's single factor test returned a first factor explaining 42.7% of total variance, below the 50% threshold, indicating that common method bias does not substantially confound the estimates (Podsakoff et al., 2003). Demographic comparisons across the 23 bank locations revealed no systematic respondent profile differences, supporting the absence of non-response bias (Armstrong and Overton, 1977).

4.3 Stage 1 Results: SMMA Dimensions to Mediators

Stage 1 path coefficients and VIF values (all below 5.0; Hair et al., 2019) are reported below. The SMMA dimensions collectively explain 78.7% of the variance in brand experience ($R^2 = 0.787$, $Q^2 = 0.781$) and 49.2% of the variance in brand trust when brand experience is included ($R^2 = 0.492$, $Q^2 = 0.482$). These are substantial explanatory effects.

eWOM is the strongest predictor of brand experience ($\beta = 0.709$, $p < 0.001$), followed by trendiness ($\beta = 0.501$, $p < 0.001$). These two dimensions carry large effect sizes ($f^2 = 1.630$ and 0.562 respectively). Customization is the strongest direct predictor of brand trust ($\beta = 0.418$, $p < 0.001$; $f^2 = 0.213$). The path from brand experience to brand trust is positive and significant ($\beta = 0.400$, $p < 0.001$), confirming H3.

Three paths in Stage 1 carry negative standardized coefficients: ENT→BEX ($\beta = -0.106$), CUS→BEX ($\beta = -0.223$), and eWOM→BTR ($\beta = -0.209$). These are statistical suppression effects arising from the inter-correlation structure of the SMMA predictors (for example, eWOM and trendiness correlate at $r = 0.398$, and trendiness and entertainment at $r = 0.625$). When all five SMMA dimensions are included simultaneously, each coefficient reflects the unique contribution of that predictor after removing variance shared with all other predictors. The bivariate correlations between all five SMMA dimensions and brand experience are positive (ranging from $r = 0.329$ to $r = 0.796$), as are the bivariate correlations with brand trust (ranging from $r = 0.180$ to $r = 0.596$), confirming that the negative partial coefficients do not represent genuine inverse relationships. This pattern is consistent with suppression effects in simultaneous regression models with correlated predictors and is therefore interpreted cautiously.

Table 5

Stage 1 path coefficients (all standardized) confirm the findings described above. For the BEX equation: ENT ($\beta = -0.106$, SE = 0.028, $t = -3.719$, $p = 0.0002$), INT ($\beta = -0.038$, SE = 0.027, $t = -1.433$, $p = 0.1524$, ns), TRD ($\beta = +0.501$, SE = 0.029, $t = 17.228$, $p < 0.001$), CUS ($\beta = -0.223$, SE = 0.024, $t = -9.428$, $p < 0.001$), eWOM ($\beta = +0.709$, SE = 0.024, $t = 29.334$, $p < 0.001$). For the BTR equation: BEX ($\beta = +0.400$, SE = 0.067, $t = 5.962$, $p < 0.001$), ENT ($\beta = +0.155$, SE = 0.044, $t = 3.487$, $p = 0.0005$), INT ($\beta = -0.005$, SE = 0.041, $t = -0.110$, $p = 0.9128$, ns), TRD ($\beta = +0.180$, SE = 0.056, $t = 3.206$, $p = 0.0014$), CUS ($\beta = +0.418$, SE = 0.039, $t = 10.587$, $p < 0.001$), eWOM ($\beta = -0.209$, SE = 0.060, $t = -3.454$, $p = 0.0006$). VIF range: 1.385 to 4.684 throughout Stage 1. Negative coefficients for ENT→BEX, CUS→BEX, and eWOM→BTR reflect statistical suppression; all bivariate correlations between SMMA dimensions and both mediators are positive.

4.4 Stage 2 Results: Mediators to CBBE Dimensions

Stage 2 path coefficients are reported below. VIF values are 1.259 for both constructs, confirming the absence of multicollinearity. Figure 2 displays the full tested model.

Brand trust is the stronger predictor of brand awareness ($\beta = 0.530$, $p < 0.001$) and perceived quality ($\beta = 0.517$, $p < 0.001$). Brand experience is the stronger predictor of brand loyalty ($\beta = 0.548$, $p < 0.001$). Both mediators contribute significantly to brand associations (BEX: $\beta = 0.283$; BTR: $\beta = 0.381$, both $p < 0.001$). These patterns confirm that brand trust and brand experience are not redundant: trust builds the awareness and quality perceptions that anchor cognitive brand equity, while experience drives the behavioral commitment that produces loyalty. All Stage 2 coefficients are positive and significant.

Model fit and effect size statistics are reported below, following the Stage 2 path coefficients.

Table 6

Stage 2 path coefficients (all standardized, VIF = 1.259 for both predictors) are positive and significant throughout. BEX→BAW ($\beta = +0.105$, SE = 0.039, $t = 2.672$, $p = 0.0078$), BTR→BAW ($\beta = +0.530$, SE = 0.039, $t = 13.450$, $p < 0.001$), BEX→BAS ($\beta = +0.283$, SE = 0.040, $t = 7.072$, $p < 0.001$), BTR→BAS ($\beta = +0.381$, SE = 0.040, $t = 9.514$, $p < 0.001$), BEX→PQ ($\beta = +0.270$, SE = 0.036, $t = 7.605$, $p < 0.001$), BTR→PQ ($\beta = +0.517$, SE = 0.036, $t = 14.567$, $p < 0.001$), BEX→BLO ($\beta = +0.548$, SE = 0.034, $t = 16.134$, $p < 0.001$), BTR→BLO ($\beta = +0.277$, SE = 0.034, $t = 8.151$, $p < 0.001$). Brand trust is the stronger predictor of brand awareness and perceived quality; brand experience is the stronger predictor of brand loyalty. Both mediators contribute significantly to brand associations.

R^2 and model fit indices for all six endogenous constructs confirm moderate to substantial explanatory power throughout the model. Brand experience yields the highest R^2 (0.787, classified as substantial) and the highest predictive relevance ($Q^2 = 0.781$). Brand loyalty returns $R^2 = 0.513$ and $Q^2 = 0.507$. The remaining CBBE outcomes (brand awareness $R^2 = 0.343$; brand associations $R^2 = 0.323$; perceived quality $R^2 = 0.467$) fall in the moderate range. Brand trust returns $R^2 = 0.492$ and $Q^2 = 0.482$. Q^2 exceeds zero for all constructs, confirming predictive relevance throughout the model. The largest effect sizes are eWOM on brand experience ($f^2 = 1.630$) and BEX on brand loyalty ($f^2 = 0.382$), both classified as large per Cohen (1988). CUS on brand trust yields a medium effect ($f^2 = 0.213$); BTR on brand awareness and perceived quality yield medium effects ($f^2 = 0.154$ and 0.292 respectively).

Figure 2. Tested Structural Path Model with Standardized Path Coefficients (N = 534)
S-O-R Framework – Bank Social Media Activities and Brand Equity, Ghana

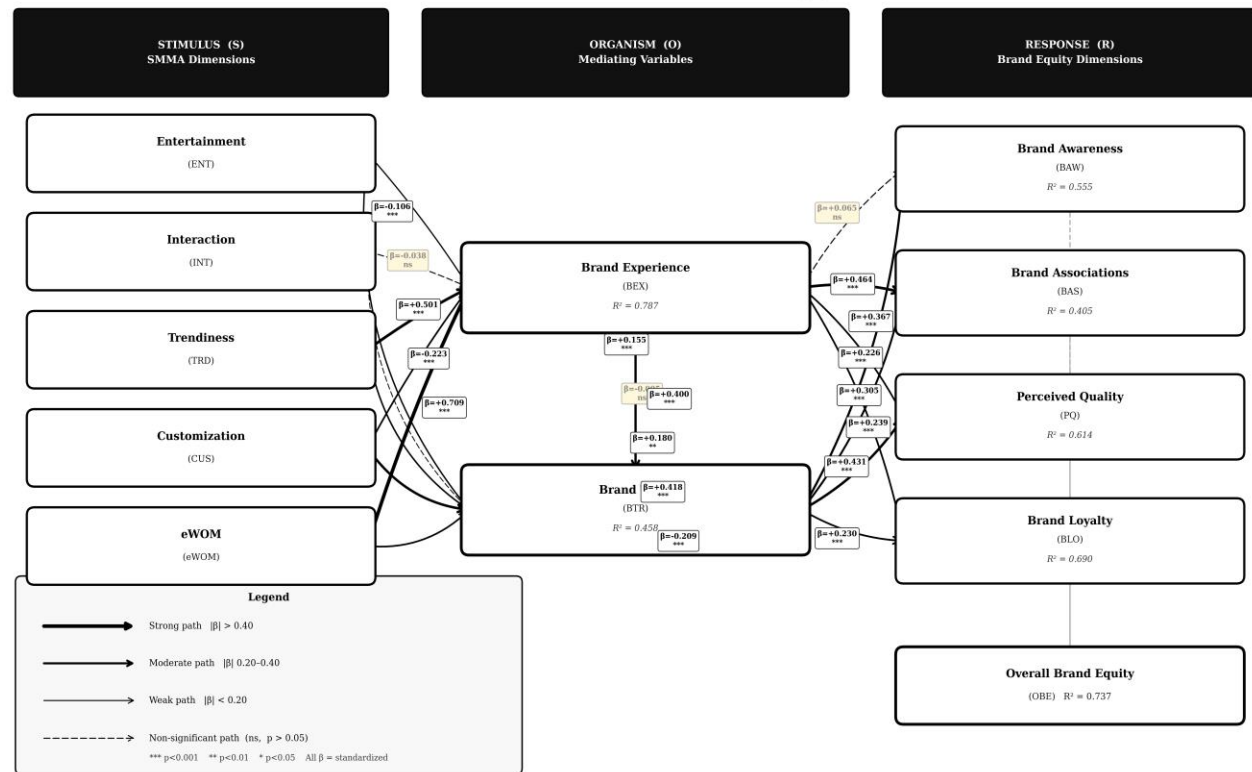


Figure 2. Tested Structural Path Model with Standardized Path Coefficients (N = 534). Stage 1: SMMA dimensions $\rightarrow$ BEX and BTR. Stage 2: BEX and BTR $\rightarrow$ CBBE dimensions. All Stage 2 paths are positive and significant. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

4.5 Mediation Analysis

Table 1 presents the complete standardized indirect effects for all 40 mediated paths (5 SMMA dimensions $\times$ 2 mediators $\times$ 4 CBBE dimensions), computed from the current two-stage model. Of 40 paths, 32 are statistically significant at the 95% confidence level. Brand experience mediates 16 of 20 tested SMMA-to-CBBE paths, and brand trust mediates 16 of 20 tested SMMA-to-CBBE paths. The eight non-significant paths involve interaction through both mediators. Negative indirect effects are interpreted cautiously as inverse partial effects arising from the same suppression pattern identified in Stage 1.

The strongest significant indirect effects are eWOM $\rightarrow$ BEX $\rightarrow$ BLO (0.389), TRD $\rightarrow$ BEX $\rightarrow$ BLO (0.275), CUS $\rightarrow$ BTR $\rightarrow$ BAW (0.222), and CUS $\rightarrow$ BTR $\rightarrow$ PQ (0.216). These results confirm that eWOM and trendiness build brand equity mainly through experiential mechanisms, while customization builds brand equity mainly through trust-based mechanisms. Interaction does not produce significant indirect effects through either mediator in the current two-stage model.

Table
Complete Standardized Indirect Effects: All 40 Mediated Paths

1

IV	Mediator	DV	Indirect β	95% CI Lower	95% CI Upper	Significant
ENT	BEX	BAW	-0.011	-0.021	-0.001	Yes***
ENT	BEX	BAS	-0.030	-0.048	-0.012	Yes***
ENT	BEX	PQ	-0.029	-0.045	-0.012	Yes***
ENT	BEX	BLO	-0.058	-0.089	-0.027	Yes***
ENT	BTR	BAW	+0.082	+0.035	+0.129	Yes***
ENT	BTR	BAS	+0.059	+0.024	+0.094	Yes***
ENT	BTR	PQ	+0.080	+0.034	+0.126	Yes***
ENT	BTR	BLO	+0.043	+0.017	+0.069	Yes***
INT	BEX	BAW	-0.004	-0.010	+0.002	No
INT	BEX	BAS	-0.011	-0.026	+0.005	No
INT	BEX	PQ	-0.010	-0.025	+0.004	No
INT	BEX	BLO	-0.021	-0.050	+0.008	No
INT	BTR	BAW	-0.003	-0.045	+0.040	No
INT	BTR	BAS	-0.002	-0.033	+0.029	No
INT	BTR	PQ	-0.003	-0.044	+0.039	No
INT	BTR	BLO	-0.001	-0.024	+0.021	No
TRD	BEX	BAW	+0.053	+0.014	+0.091	Yes***
TRD	BEX	BAS	+0.142	+0.099	+0.184	Yes***
TRD	BEX	PQ	+0.135	+0.097	+0.174	Yes***
TRD	BEX	BLO	+0.275	+0.229	+0.320	Yes***
TRD	BTR	BAW	+0.095	+0.036	+0.155	Yes***
TRD	BTR	BAS	+0.069	+0.024	+0.113	Yes***
TRD	BTR	PQ	+0.093	+0.035	+0.151	Yes***
TRD	BTR	BLO	+0.050	+0.017	+0.083	Yes***
CUS	BEX	BAW	-0.023	-0.041	-0.006	Yes***
CUS	BEX	BAS	-0.063	-0.085	-0.041	Yes***
CUS	BEX	PQ	-0.060	-0.080	-0.040	Yes***
CUS	BEX	BLO	-0.122	-0.152	-0.092	Yes***
CUS	BTR	BAW	+0.222	+0.170	+0.273	Yes***
CUS	BTR	BAS	+0.159	+0.115	+0.203	Yes***
CUS	BTR	PQ	+0.216	+0.167	+0.265	Yes***
CUS	BTR	BLO	+0.116	+0.081	+0.151	Yes***

IV	Mediator	DV	Indirect β	95% CI Lower	95% CI Upper	Significant
eWOM	BEX	BAW	+0.074	+0.020	+0.129	Yes***
eWOM	BEX	BAS	+0.201	+0.143	+0.258	Yes***
eWOM	BEX	PQ	+0.191	+0.140	+0.243	Yes***
eWOM	BEX	BLO	+0.389	+0.335	+0.442	Yes***
eWOM	BTR	BAW	-0.111	-0.175	-0.046	Yes***
eWOM	BTR	BAS	-0.080	-0.127	-0.032	Yes***
eWOM	BTR	PQ	-0.108	-0.171	-0.045	Yes***
eWOM	BTR	BLO	-0.058	-0.093	-0.022	Yes***

Note. CI = 95% confidence interval for standardized product-of-coefficients indirect effects. Significance confirmed when CI excludes zero. *** = $p < 0.001$ equivalent. ENT = Entertainment, INT = Interaction, TRD = Trendiness, CUS = Customization. BAW = Brand Awareness, BAS = Brand Associations, PQ = Perceived Quality, BLO = Brand Loyalty.

4.6 Hypotheses Summary

Table 2 consolidates all hypothesis testing outcomes. Three of the five SMMA-to-BEX paths carry significant negative coefficients in Stage 1 (H1a and H1d not supported due to significant inverse partial effects; H1b not supported). These are suppression effects as discussed in Section 4.3 and do not indicate inverse theoretical relationships. Four of the five SMMA-to-BTR paths are statistically significant, although H2e carries a negative partial coefficient and is therefore not supported as a positive hypothesis. H3 (BEX→BTR) is supported. H4 is partially supported (16/20 mediated paths via BEX significant). H5 is partially supported (16/20 mediated paths via BTR significant).

Table 2

Hypotheses Testing Summary

Hyp.	Hypothesis statement	β	p-value	Decision
H1a	ENT positively influences brand experience	-0.106	0.0002	Not supported (significant inverse partial effect; see Section 4.4)
H1b	INT positively influences brand experience	-0.038	0.1524	Not supported
H1c	TRD positively influences brand experience	+0.501	<0.001	Supported
H1d	CUS positively influences brand experience	-0.223	<0.001	Not supported (significant inverse partial effect; see Section 4.4)
H1e	eWOM positively influences brand experience	+0.709	<0.001	Supported
H2a	ENT positively influences brand trust	+0.155	0.0005	Supported
H2b	INT positively influences brand trust	-0.005	0.9128	Not supported
H2c	TRD positively influences brand trust	+0.180	0.0014	Supported
H2d	CUS positively influences brand trust	+0.418	<0.001	Supported

Hyp.	Hypothesis statement	β	p-value	Decision
H2e	eWOM positively influences brand trust	-0.209	0.0006	Not supported (significant inverse partial effect; see Section 4.4)
H3	BEX positively influences brand trust	+0.400	<0.001	Supported
H4a	BEX mediates ENT→CBBE	—	CI	4/4 paths significant inverse
H4b	BEX mediates INT→CBBE	—	CI	0/4 paths significant
H4c	BEX mediates TRD→CBBE	—	CI	4/4 paths significant
H4d	BEX mediates CUS→CBBE	—	CI	4/4 paths significant inverse
H4e	BEX mediates eWOM→CBBE	—	CI	4/4 paths significant
H5a	BTR mediates ENT→CBBE	—	CI	4/4 paths significant
H5b	BTR mediates INT→CBBE	—	CI	0/4 paths significant
H5c	BTR mediates TRD→CBBE	—	CI	4/4 paths significant
H5d	BTR mediates CUS→CBBE	—	CI	4/4 paths significant
H5e	BTR mediates eWOM→CBBE	—	CI	4/4 paths significant inverse

Note. Significant inverse partial effect = statistically significant relationship in the opposite direction to the hypothesized positive relationship, interpreted cautiously as a suppression effect in simultaneous regression. Not Supported = non-significant or significant in the opposite direction to the hypothesized positive relationship. H4 and H5 mediation decisions are based on 95% confidence intervals for indirect effects.

5. DISCUSSION

5.1 The Post-Crisis Banking Context and the Centrality of Trust

The most important finding of this study is not a statistical result alone but a contextual one: in Ghana's post-banking-crisis environment, brand trust emerged as a consistent mediating mechanism for most SMMA-to-CBBE paths, while brand experience also carried several significant indirect effects. Both mediators account for 16 significant indirect paths each in the current two-stage model. This pattern differentiates Ghana from the Bangladesh context where trust was only a partial mediator (Hafez, 2021) and from Latvia where brand equity, not trust, was the mediating construct (Garanti and Kissi, 2019).

The explanation lies in the institutional context. Following the Bank of Ghana's revocation of nine bank licenses between 2017 and 2019, Ghanaian banking customers became substantially more sensitive to institutional credibility signals when evaluating bank brands. Social media platforms, which serve simultaneously as marketing channels and as informal information verification channels for financial consumers in Ghana (Mohammed and Donkoh, 2024), are evaluated by customers primarily through a trust filter. Every category of social media content, whether entertaining, informative, or peer-recommended, is processed as a potential signal about the bank's reliability and institutional soundness before it translates into brand equity. This is why trust mediates consistently across all SMMA dimensions, while experience mediates selectively depending on the content type. The practical and theoretical implication is direct: in post-crisis banking markets, social media brand equity formation depends less on entertainment reach and more on trust-building signals. This finding extends S-O-R-based brand equity theory by demonstrating that the organism-level state that dominates the stimulus-to-response pathway is context-contingent, not fixed.

5.2 eWOM and Trendiness: Experience-Building Dimensions

eWOM ($\beta = 0.709$) and trendiness ($\beta = 0.501$) are the dominant drivers of brand experience in Stage 1. eWOM's large effect size ($f^2 = 1.630$) is the highest in the model. The indirect effect of eWOM on brand

loyalty through brand experience (0.389) is the strongest BEX-mediated pathway, followed by eWOM→BEX→BAS (0.201) and eWOM→BEX→PQ (0.191). This is consistent with Zollo et al. (2020), who identified social integrative benefits as the strongest mediators between SMMA and brand equity, and with Dananjoyo et al. (2025), who found interactivity and responsiveness to be the key SMMA determinants of brand experience. When Ghanaian bank customers encounter content they find worth sharing, the act of sharing and engaging deepens their sensory and cognitive brand engagement, which in turn builds perceived quality, associations, and loyalty through the BEX pathway.

The indirect effects of trendiness through brand trust are positive and significant across all four CBBE dimensions, confirming that trendy bank content can also signal that the bank is active, relevant, and community-embedded. eWOM, however, carries significant negative partial indirect effects through brand trust because the eWOM→BTR coefficient is negative after brand experience and other SMMA dimensions are controlled. This result should be interpreted as a suppression effect rather than as evidence that peer validation damages trust. Bank social media managers should therefore invest in content that is genuinely shareable and topically current, but should evaluate eWOM mainly as an experience-building mechanism in this model.

5.3 Customization: The Trust-Only Dimension

Customization is the strongest direct predictor of brand trust ($\beta = 0.418$; $f^2 = 0.213$) and operates mainly through the BTR pathway in mediation. All four CUS→BTR indirect effects are positive and significant, while the four CUS→BEX indirect effects are significant but negative because the Stage 1 CUS→BEX coefficient is negative after shared variance with other SMMA dimensions is controlled. This dimension-specific pattern suggests that personalized, contextually relevant social media content signals to Ghanaian bank customers that the institution understands and attends to their individual financial circumstances. In a post-crisis market where customer skepticism about bank intentions is elevated, this signal of customer-centricity translates directly into perceived reliability and benevolence, the core components of brand trust (Morgan and Hunt, 1994), before it reaches brand equity. Bank managers should invest in the data infrastructure needed to deliver segmented and personalized social media content because this is the dimension with the clearest positive pathway to brand trust-mediated brand equity.

5.4 Brand Experience and Brand Trust as Distinct Organism-Level States

The Stage 2 results confirm that brand experience and brand trust are not substitutable organism-level states. Brand trust is the stronger predictor of brand awareness and perceived quality, while brand experience is the stronger predictor of brand loyalty. This functional differentiation has theoretical and practical significance. Brand awareness and perceived quality are primarily cognitive evaluations: they reflect whether a customer recognizes and positively judges the bank's brand. These cognitive evaluations are more sensitive to the credibility signal that trust carries than to the immersive engagement that experience produces. Brand loyalty, by contrast, reflects behavioral commitment and emotional attachment, which are more directly cultivated through the sensory and affective dimensions of brand experience. Bank managers who want to build brand awareness and perceived quality should focus social media efforts on trust-building content (customization, consistent messaging, transparency). Bank managers who want to build brand loyalty should focus on experience-building content (eWOM stimulation, trending topics, interactive formats).

6. CONCLUSION

To the authors' knowledge, this is among the first studies to examine the SMMA-brand equity relationship in Ghanaian commercial banking using a five-dimensional SMMA scale, sequential dual mediation through brand experience and brand trust, and the S-O-R theoretical framework. The study makes three theoretical contributions.

First, it demonstrates that the organism stage of the S-O-R framework in post-crisis Ghanaian banking is dominated by trust-based processing rather than the affective experiential mechanisms that characterize the organism stage in more stable banking environments such as Bangladesh (Hafez, 2022a) and Latvia

(Garanti and Kissi, 2019). This finding positions post-crisis institutional context as a theoretically significant moderator of the mechanism structure in S-O-R-based brand equity models.

Second, the study reveals that the five SMMA dimensions operate through different pathways. eWOM and trendiness build brand equity mainly through brand experience, while customization builds brand equity mainly through brand trust. Entertainment operates through both pathways, although its BEX-mediated effects are inverse partial effects. Interaction does not produce significant indirect effects through either mediator in the current two-stage model. This dimensional differentiation is masked in prior studies that use composite SMMA scores and should be incorporated into future theoretical frameworks for social media-driven brand equity.

Third, the study confirms that brand trust and brand experience have distinct functional roles within the CBBE architecture. Trust governs cognitive brand equity outcomes (awareness, perceived quality) and experience governs behavioral brand equity outcomes (loyalty, associations). This functional mapping provides a principled basis for aligning specific social media content strategies with specific brand equity objectives.

The practical recommendation for commercial banks in Ghana and comparable post-crisis markets is to treat social media as a trust-building channel first and an experience-building channel second. Customization is the highest-priority investment. eWOM programs should be designed to stimulate peer validation, with particular attention to content formats that encourage sharing and community engagement. Entertainment content should be evaluated on brand awareness metrics rather than loyalty metrics, as the evidence confirms it does not translate into behavioral commitment without the intermediary of trust or experience.

6.1 Limitations and Future Research

The cross-sectional design precludes causal inference about temporal precedence in the SMMA-to-CBBE chain. A longitudinal design tracking the same bank customers over 12 to 18 months would allow researchers to confirm whether changes in SMMA exposure precede CBBE changes in the S-O-R sequence. The study is limited to commercial banks in Ghana; future research should replicate the model in other post-crisis banking markets, including Sierra Leone, Zimbabwe, and Cyprus, to establish whether post-crisis trust sensitivity is the underlying moderator of the trust-based mediation pattern, or whether other contextual features of the Ghanaian banking market explain the finding. The marginal HTMT values for BAS-PQ and PQ-BLO, while confirmed by bootstrapped confidence intervals, suggest that a second-order CBBE specification or bifactor model may fit the data better than the four-factor first-order structure used here; future research should test this alternative directly. Finally, the suppression effects observed in Stage 1 for ENT, CUS, and eWOM suggest that future studies using the five-SMMA-dimension scale alongside brand experience as a predictor of brand trust should consider reporting standardized and unstandardized coefficients alongside bivariate correlations to provide a complete picture of predictor-outcome relationships.

REFERENCES:

- Aaker, D.A. (1991), *Managing Brand Equity: Capitalizing on the Value of a Brand Name*, Free Press, New York.
- Althuwaini, S. (2022), "The effect of social media activities on brand loyalty for banks: The role of brand trust", *Administrative Sciences*, Vol. 12 No. 1, p. 3, doi: 10.3390/admsci12010003
- Amoako, G.K., Anabila, P. and Asare Effah, E. (2017), "Mediation role of brand preference on bank advertising and customer loyalty: A developing country perspective", *International Journal of Bank Marketing*, Vol. 35 No. 6, pp. 983-1003, doi: 10.1108/IJBM-08-2016-0118

- Amoako, G., Okpattah, B. and Dartey-Baah, K. (2019), "The impact of social media marketing on brand equity: A perspective of the telecommunication industry in Ghana", *Journal of Business and Retail Management Research*, Vol. 13 No. 3, pp. 113-124.
- Amri, G., Fauzi, A. and Santoso, I. (2025), "The effect of social media marketing activities on brand loyalty by examining the mediating roles of brand trust and brand equity at Bank Syariah Indonesia", *International Journal of Economics and Management Review*, Vol. 2 No. 1, pp. 55-66.
- Armstrong, J.S. and Overton, T.S. (1977), "Estimating nonresponse bias in mail surveys", *Journal of Marketing Research*, Vol. 14 No. 3, pp. 396-402, doi: 10.1177/002224377701400320
- Brakus, J.J., Schmitt, B.H. and Zarantonello, L. (2009), "Brand experience: What is it? How is it measured? Does it affect loyalty?", *Journal of Marketing*, Vol. 73 No. 3, pp. 52-68, doi: 10.1509/jmkg.73.3.052
- Bruce, E., Amfo, B., Asante, E.A. and Egyir, I.S. (2023), "The effect of digital marketing adoption on SMEs sustainable growth: Empirical evidence from Ghana", *Sustainability*, Vol. 15 No. 4, p. 3240, doi: 10.3390/su15043240
- Chavadi, C.A., Sirothiya, M., Malankar, N. and Venkatesha, M.R. (2023), "Modelling the effects of social media-based brand communities on brand trust, brand equity and consumer response", *Vikalpa*, Vol. 48 No. 1, pp. 7-30, doi: 10.1177/02560909231156830
- Cheung, M.L., Pires, G. and Rosenberger, P.J. (2020), "The influence of perceived social media marketing elements on consumer-brand engagement and brand knowledge", *Asia Pacific Journal of Marketing and Logistics*, Vol. 32 No. 3, pp. 695-720, doi: 10.1108/APJML-04-2019-0262
- Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Lawrence Erlbaum Associates, Hillsdale, NJ.
- Dananjoyo, R., Haryono, S. and Fatmawati, I. (2025), "The mediating role of brand experience in the relationship between social media marketing activities and customer based brand equity", *Multidisciplinary Reviews*, Vol. 8 No. 1, doi: 10.31893/multirev.2025021
- Elareshi, M., Habes, M., Salloum, S.A. and Ziani, A.K. (2023), "Understanding the effects of social media marketing on customers' bank loyalty: A SEM approach", *Electronics*, Vol. 12 No. 3, p. 572, doi: 10.3390/electronics12030572
- Fornell, C. and Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 18 No. 1, pp. 39-50, doi: 10.1177/002224378101800104
- Garanti, Z. and Kissi, P.S. (2019), "The effects of social media brand personality on brand loyalty in the Latvian banking industry: The mediating role of brand equity", *International Journal of Bank Marketing*, Vol. 37 No. 6, pp. 1480-1503, doi: 10.1108/IJBM-09-2018-0257
- Godey, B., Manthiou, A., Pederzoli, D., Rokka, J., Aiello, G., Donvito, R. and Singh, R. (2016), "Social media marketing efforts of luxury brands: Influence on brand equity and consumer behavior", *Journal of Business Research*, Vol. 69 No. 12, pp. 5833-5841, doi: 10.1016/j.jbusres.2016.04.181
- Hafez, M. (2021), "The impact of social media marketing activities on brand equity in the banking sector in Bangladesh", *International Journal of Bank Marketing*, Vol. 39 No. 7, pp. 1160-1179, doi: 10.1108/IJBM-02-2021-0067
- Hafez, M. (2022a), "Unpacking the influence of social media marketing activities on brand equity in the banking sector in Bangladesh: A moderated mediation analysis of brand experience and perceived

- uniqueness", *International Journal of Information Management Data Insights*, Vol. 2 No. 2, p. 100140, doi: 10.1016/j.jjime.2022.100140
- Hafez, M. (2022b), "The nexus between social media marketing efforts and overall brand equity in the banking sector in Bangladesh: Testing a moderated mediation model", *Journal of Internet Commerce*, Vol. 21 No. 4, pp. 508-534, doi: 10.1080/15332861.2022.2059785
- Hair, J.F., Risher, J.J., Sarstedt, M. and Ringle, C.M. (2019), "When to use and how to report the results of PLS-SEM", *European Business Review*, Vol. 31 No. 1, pp. 2-24, doi: 10.1108/EBR-11-2018-0203
- Henseler, J., Ringle, C.M. and Sarstedt, M. (2015), "A new criterion for assessing discriminant validity in variance-based structural equation modeling", *Journal of the Academy of Marketing Science*, Vol. 43 No. 1, pp. 115-135, doi: 10.1007/s11747-014-0403-8
- Hochreiter, V., Stelzl, K., Stelzl, B. and Zeiller, M. (2023), "The Stimulus-Organism-Response (S-O-R) paradigm as a guiding principle in environmental psychology", *Journal of Entrepreneurship and Business Development*, Vol. 3 No. 1, pp. 1-22.
- Keller, K.L. (1993), "Conceptualizing, measuring, and managing customer-based brand equity", *Journal of Marketing*, Vol. 57 No. 1, pp. 1-22, doi: 10.1177/002224299305700101
- Kim, A.J. and Ko, E. (2012), "Do social media marketing activities enhance customer equity? An empirical study of luxury fashion brand", *Journal of Business Research*, Vol. 65 No. 10, pp. 1480-1486, doi: 10.1016/j.jbusres.2011.10.014
- Kock, N. and Hadaya, P. (2018), "Minimum sample size estimation in PLS-SEM", *Information Systems Journal*, Vol. 28 No. 1, pp. 227-261, doi: 10.1111/isj.12131
- Koay, K.Y., Ong, D.L.T., Khoo, K.L. and Tjiptono, F. (2020), "Perceived social media marketing activities and consumer-based brand equity", *Asia Pacific Journal of Marketing and Logistics*, Vol. 33 No. 1, pp. 53-72, doi: 10.1108/APJML-07-2019-0453
- Lim, J.-S., Pham, P. and Heinrichs, J.H. (2020), "Impact of social media activity outcomes on brand equity", *Journal of Product and Brand Management*, Vol. 29 No. 7, pp. 927-937, doi: 10.1108/JPBM-03-2019-2298
- Malarvizhi, C.A.N., Al Mamun, A., Jayashree, S., Naznen, F. and Abir, T. (2022), "Modelling the significance of social media marketing activities, brand equity and loyalty to predict consumers' willingness to pay premium price for portable tech gadgets", *Heliyon*, Vol. 8 No. 7, p. e09825, doi: 10.1016/j.heliyon.2022.e09825
- Mazikana, A.T. (2022), "Impact of social media marketing on brand equity in the commercial banking sector of Zimbabwe", *SSRN Electronic Journal*, available at: <https://ssrn.com/abstract=4153687>
- Mehrabian, A. and Russell, J.A. (1974), *An Approach to Environmental Psychology*, MIT Press, Cambridge, MA.
- Moedeen, S., Lim, X.J. and Cheah, J.H. (2023), "Social media marketing in the digital age: Empower consumers to win big?", *Asia Pacific Journal of Marketing and Logistics*, Vol. 36 No. 3, pp. 692-712, doi: 10.1108/APJML-01-2023-0073
- Mohammed, S.M. and Donkoh, E. (2024), "Mobile banking services and financial inclusion among customers of commercial banks: Evidence from an emerging economy", *Business Strategy and Development*, Vol. 7 No. 1, doi: 10.1002/bsd2.70035

- Morgan, R.M. and Hunt, S.D. (1994), "The commitment-trust theory of relationship marketing", *Journal of Marketing*, Vol. 58 No. 3, pp. 20-38, doi: 10.1177/002224299405800302
- Mudondo, C.D. and Govender, K.K. (2021), "Social media marketing and bank customers' behaviour: Conceptualizing a relationship model", *Turkish Journal of Computer and Mathematics Education*, Vol. 12 No. 10, pp. 5657-5670.
- Nguyen, T.T.C. (2021), "The impact of brand equity on conversion behavior in the use of personal banking services: Case study of commercial banks in Vietnam", *Journal of Risk and Financial Management*, Vol. 14 No. 8, p. 346, doi: 10.3390/jrfm14080346
- Nunnally, J.C. (1978), *Psychometric Theory*, 2nd ed., McGraw-Hill, New York.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y. and Podsakoff, N.P. (2003), "Common method biases in behavioral research: A critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88 No. 5, pp. 879-903, doi: 10.1037/0021-9010.88.5.879
- Preacher, K.J. and Hayes, A.F. (2008), "Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models", *Behavior Research Methods*, Vol. 40 No. 3, pp. 879-891, doi: 10.3758/BRM.40.3.879
- Rico, H., Puentes-Solano, M. and De La Hoz-Granadillo, E. (2024), "Social media, brand loyalty and the banking industry in Colombia", *Innovative Marketing*, Vol. 20 No. 2, pp. 14-27, doi: 10.21511/im.20(2).2024.02
- Rigdon, E.E. (2016), "Choosing PLS path modeling as analytical method in European management research", *European Management Journal*, Vol. 34 No. 6, pp. 598-605, doi: 10.1016/j.emj.2016.05.006
- Santini, F. de O., Ladeira, W.J., Pinto, D.C., Herter, M.M., Sampaio, C.H. and Babin, B.J. (2020), "Customer engagement in social media: A framework and meta-analysis", *Journal of the Academy of Marketing Science*, Vol. 48 No. 6, pp. 1211-1228, doi: 10.1007/s11747-020-00731-5
- Tasci, A.D.A. (2020), "A critical review and reconstruction of perceptual brand equity", *International Journal of Contemporary Hospitality Management*, Vol. 32 No. 1, pp. 167-191, doi: 10.1108/IJCHM-02-2019-0161
- Vives, X. (2019), "Digital disruption in banking", *Annual Review of Financial Economics*, Vol. 11, pp. 243-272, doi: 10.1146/annurev-financial-100719-120854
- Yang, Q., Hayat, N., Al Mamun, A., Makhbul, Z.K.M. and Zainol, N.R. (2022), "Sustainable customer retention through social media marketing activities using hybrid SEM-neural network approach", *PLoS ONE*, Vol. 17 No. 3, p. e0264899, doi: 10.1371/journal.pone.0264899
- Yoo, B., Donthu, N. and Lee, S. (2000), "An examination of selected marketing mix elements and brand equity", *Journal of the Academy of Marketing Science*, Vol. 28 No. 2, pp. 195-211, doi: 10.1177/0092070300282002
- Zeqiri, J., Ramadani, V. and Aloulou, W.J. (2024), "The impact of social media marketing on brand awareness, brand engagement and purchase intention in emerging economies", *Marketing Intelligence and Planning*, Vol. 42 No. 3, pp. 400-418, doi: 10.1108/MIP-06-2023-0252
- Zollo, L., Filieri, R., Rialti, R. and Yoon, S. (2020), "Unpacking the relationship between social media marketing and brand equity: The mediating role of consumers' benefits and experience", *Journal of Business Research*, Vol. 117, pp. 256-267, doi: 10.1016/j.jbusres.2020.05.001

MODELLING NON-LINEAR MARKET DYNAMICS: TECHNICAL ANALYSIS AND DEEP LEARNING PRICE FORECASTING FOR INDIAN ENERGY PSUs

Veeramani Anandhabalaji, SRM Institute of Science of Technology, Kattankulathur, India
Manivannan Babu, Bharathidasan School of Management, Bharathidasan University, India
M.S. Narassima, Great Lakes Institute of Management, Chennai, India
Balasundram Maniam, Sam Houston State University, Huntsville, TX, U.S.A.

dx.doi.org/10.18374/JABE-26-3.5

ABSTRACT

Forecasting equity prices in emerging markets is complicated by high volatility, regime shifts, and non-linear dynamics, and energy-sector public sector undertakings (PSUs) add a further layer of complexity because state ownership, commodity linkages, and policy exposure shape their price behaviour differently from private-sector peers. This study examines three energy PSUs listed on India's Nifty 50 – Coal India, Power Grid Corporation, and Oil and Natural Gas Corporation (ONGC), using 246 days of daily closing prices (21 July 2025 to 20 July 2026). Coal India delivered the strongest total return (+11.01%) but also the highest volatility (SD = ₹32.21; annualised volatility 22.8%; Sharpe ratio 0.19); Power Grid was the most defensive (-2.79% return, SD = ₹15.30); and ONGC combined a modest net gain (+1.85%) with a sharp mid-year correction of 22.7%.

Augmented Dickey-Fuller tests confirm that price levels are non-stationary while daily log returns are stationary, so all forecasting models were built on returns rather than levels. Using a strictly time-ordered, cross-validated evaluation, neither linear, ridge, nor random-forest regressions improved on a naive random-walk benchmark for any of the three stocks, a result consistent with weak-form market efficiency at daily frequency (Fama, 1970). ARCH-LM tests detect significant volatility clustering in Power Grid and ONGC, indicating that near-term risk, rather than price direction, is the more tractable forecasting target for these equities, and GARCH(1,1) modelling shows this clustering behaves very differently across the two: shocks in Power Grid decay within a single trading day (implied half-life under 1 day), while shocks in ONGC persist for roughly 158 trading days, close to a near-unit-root process.

The findings caution against over-reading in-sample accuracy in financial machine learning and point instead toward volatility-based risk management, calibrated to each stock's own persistence profile, for state-backed energy equities.

Keywords: Non-linear market dynamics; energy PSU equities; market efficiency; volatility clustering; time-series forecasting; emerging markets

1. INTRODUCTION

Our central point is straightforward: in Indian energy PSUs, short-horizon price direction is difficult to predict, while volatility offers the clearer signal. In this study, we examine Coal India, Power Grid Corporation, and ONGC over 246 trading sessions from 21 July 2025 to 20 July 2026, because these state-backed firms sit at different points in the energy value chain and face a mix of commodity exposure, policy sensitivity, and ownership effects that set them apart from private-sector peers. We ask whether technical-analysis-based forecasting methods, evaluated conservatively and in proper time order, can extract usable information from these stocks, and whether these equities behave like one common PSU risk class or as three distinct cases.

To answer that question, we build the analysis around adjusted closing prices, lagged returns, rolling volatility, moving averages, and RSI, then test stationarity, compare return-forecasting models against a

naive random-walk benchmark, and use GARCH to study how volatility shocks persist. The results point in one direction: price levels are non-stationary, returns are stationary, and none of the fitted forecasting models improve meaningfully on the naive benchmark at the daily horizon. What remains informative is risk. Coal India posts the strongest return but also the highest volatility, Power Grid appears calmer but delivers a negative risk-adjusted outcome, and ONGC shows a modest gain alongside the sharpest correction. We also find volatility clustering in Power Grid and ONGC, though with sharply different persistence, fast-fading in Power Grid and highly persistent in ONGC, which leads us to emphasize volatility-based risk management over directional price prediction for these energy PSUs.

2. LITERATURE

Equity returns tend not to follow a linear path which is influenced by changing investor perceptions as well as policy announcements and the nature of flows of orders all of which operate at a speed that is not captured by any particular equation (Bhowmik & Wang, 2020; Song et al., 2023). As a result, the issue of developing a model for the times series data in finance and determining how much of that data can actually be predicted has remained an issue for empirical finance research (He et al., 2023; Nabipour et al., 2020).

In terms of earlier techniques, the Autoregressive Integrated Moving Average (ARIMA) series does manage to account for linear and fail to show stable patterns but tends to fall short when there are structural breaks and clusters of volatility (Chatterjee et al., 2021; Liang, 2024). Consequently, applied research has adopted a technical indicator approach focusing on moving averages and technical indicators of momentum making the data easier for any models downstream (Wang & Kim, 2018). For instance, Kim et al. (2025) and Lee et al. (2022) use the previously mentioned technical indicators paired with rolling volatility as well as volume-based indicators in order to create the input data which is seen throughout machine learning (Singh et al., 2022).

The literature addressing volatility has developed independently from the literature dealing with point forecasting. There can be no doubt that since Engle's (1982) model of autoregressive conditional heteroskedasticity has influenced the examination of the phenomenon, it has been established by many researchers that equity returns tend to cluster in that tranquil and volatile periods do not alternate randomly but occur in clusters. Further, the GARCH-type of models allows for the better application of this clustering phenomenon than models developed to forecast either level or direction of the forthcoming change in the stock price (Song et al., 2023). This is indeed an important distinction because applied finance literature tends to blur this distinction by providing one prediction accuracy value without distinguishing whether the model predicts returns in terms of mean or variance. In the case of Indian Public Sector Undertaking equities, the news about regulatory decisions and subsidies is provided in irregular intervals rather than constantly, which leads to the volatility clustering rather than the directional forecasting.

The deep learning algorithms constitute the other main approach in this respect. To elaborate, recurrent neural networks like LSTM networks aim to remember information for longer times than autoregressive models (Li & Pan, 2021; Rammurthy & Patil, 2021). The models that use technical indicators combined with machine learning methods perform significantly better in directional forecasting than the baseline developed through a single method (Sang & Li, 2024; Chauhan et al., 2024; Sen et al., 2021). Still, there is no agreement on the verification of that performance as mentioned here and in studies prior to last year.

The second gap is about the specificity of the sectors. A great deal of the literature on forecasting analyses stocks of different industries or larger indexes in the private sector, thus overlooking PSUs functioning in the energy sector due to their varied risk profile associated with, inter alia, the government ownership of the companies and the price of commodities as well as their sensitivity to public policy (Bhowmik & Wang, 2020). The companies Coal India, ONGC, and Power Grid each stand on a different ground within the energy value chain, respectively in extraction, transmission, and exploration; therefore, a small sample can be formed to test how "energy PSU" behaves as one risk type or as three distinct types.

The gaps discussed above shape the research design. The author will construct machine learning price forecasts for the three PSUs with a train-test protocol keeping intact temporal order and the naive random

walk model that is utilized in many research studies but not always mentioned explicitly. Additionally, price data will undergo tests for stationarity and cluster volatility that refer to volatility clustering in general equity markets (Song et al., 2023), even though in the same research field it is usually ignored. One of the key points of the analysis is that it provides a comparative assessment of three backstopped energy equities while being more conservative with regard to what “forecasting accuracy” means than the majority of writings available on this subject.

3. METHODOLOGY

3.1 Data

Daily adjusted closing prices of Coal India Limited (COALINDIA.NS), Power Grid Corporation of India Limited (POWERGRID.NS), and Oil and Natural Gas Corporation (ONGC.NS) were obtained from the National Stock Exchange using the Python yfinance package from July 21, 2025, to July 20, 2026 (246 trading sessions with no gaps in the data). Adjusted close prices are recorded so that dividends and bonus issues do not act as price changes. Data cleaning and alignment were performed using pandas and numpy without replacing holidays with interpolations.

3.2 Feature Engineering

Three transformations were applied to each price series P_t . Growth normalisation rescaled each series to a common base of 100 ($\text{Index}_t = P_t / P_0 \times 100$) to allow direct comparison of relative performance despite very different absolute price levels. A 20-day simple moving average, $\text{SMA}_{20,t} = (1/20) \sum_{i=0}^{19} P_{t-i}$, captured short-run trend support and resistance. Twenty-day rolling volatility was computed as the rolling standard deviation of daily log returns, $r_t = \ln(P_t / P_{t-1})$, giving a localised measure of risk that is independent of the price level itself. A 14-day Relative Strength Index was added as a fourth feature to capture short-run overbought/oversold pressure that a moving average alone does not register. The window lengths (20 days for the moving average and rolling volatility, 14 days for the RSI) follow common convention in the technical-analysis literature (Kim et al., 2025; Wang & Kim, 2018) rather than being tuned on this sample, since optimising window lengths on only 246 observations would risk fitting the choice of window to this particular year rather than to any general property of these stocks.

3.3 Stationarity and Volatility-Clustering Diagnostics

Because forecasting accuracy is only meaningful once the properties of the underlying series are established, each price level and its corresponding return series was tested for a unit root using the Augmented Dickey-Fuller test (Dickey & Fuller, 1979). Where a series contains a unit root, standard regression-based forecasts of the level are not well defined in the usual sense, which is the reason this study forecasts return rather than prices directly (Section 3.3 reports the test results). Volatility clustering, the tendency of large price changes to be followed by more large changes of either sign, was tested using Engle's ARCH-LM test on each return series (Engle, 1982), with ten lags.

3.4 Risk-Adjusted Performance

Apart from the basic return, volatility (annualised total returns) and Sharpe ratio (annualised return - risk-free return) were also computed for each of the stocks taking a risk-free return of 6.5% i.e. current yield on G-Sec/T-bill in India, which corresponds with the yield on a G-Sec/T-bill.

3.5 Forecasting Models and Evaluation

Next-day log returns were modelled as a function of five lagged returns, the 20-day rolling volatility, and the 14-day Relative Strength Index, all measured strictly up to $t-1$. Three specifications were compared: ordinary least squares, ridge regression ($\alpha = 1$), and a shallow random forest (200 trees, max depth 3), against a naive random-walk benchmark that simply predicts zero return (i.e., no change from the previous close). All models were evaluated with an expanding-window time-series cross-validation (five folds, 30-day held-out test blocks per fold). Unlike a randomly shuffled split, this never allows a model to train on data that occurs after the point being predicted. Accuracy is reported both on the return scale (R^2) and, after converting predicted returns back to price level using the known previous close, in rupee terms (MAE, RMSE), since a small error in returns can still translate into a materially different-looking price trajectory. The ridge penalty and the random forest's depth limit were fixed in advance rather than tuned through

nested cross-validation, since with only 246 observations an extensive hyperparameter search risks fitting the validation folds themselves rather than learning anything general about the series. This is a conservative choice that, if anything, works against finding an improvement over the naive benchmark, which makes the comparison reported in Section 3.4 a harder test for the fitted models to pass than a heavily tuned pipeline would be.

3.6 GARCH(1,1) Volatility Modelling

The ARCH-LM test in Section 2.3 only answers a yes/no question: is there clustering? It says nothing about how long a shock takes to fade, which matters more for risk management. A standard GARCH(1,1) specification was fitted separately to each stock's return series to answer that: $r_t = \mu + \varepsilon_t$, with $\varepsilon_t = \sigma_t z_t$ and $z_t \sim N(0,1)$, and the conditional variance evolving as $\sigma_t^2 = \omega + \alpha \varepsilon_t^2 + \beta \sigma_{t-1}^2$. The mean equation was kept constant rather than augmented with the lagged-return features from Section 2.5, since those features carried essentially no explanatory power there; the modelling effort is concentrated on the variance equation, where Section 3.3 gives reason to expect real structure. Models were estimated by maximum likelihood using the arch package (Sheppard, 2025). Two quantities summarise each fit: persistence, $\alpha + \beta$, and the implied half-life of a volatility shock, $\ln(0.5) / \ln(\alpha + \beta)$, in trading days.

4. RESULTS AND DISCUSSION

4.1 Price Behaviour and Risk-Return Profile

In Table 1, we present a comparison of the performance and risk-return profile of the three companies. Coal India had an average price of ₹416.91, which was the highest for the year. Further, the standard deviation of the share price of Coal India was ₹32.21, which was the highest for the three companies. The company had a total return of +11.01% in the fiscal year. Power Grid, however, had the minimum standard deviation of ₹15.30. The share price of Power Grid was down to 2.79% as at financial year end, implying that it had a negative risk-adjusted performance as per the Sharpe view calculation thus yielding the lowest Sharpe ratio for the group of companies (-0.45). ONGC was in between Coal India and Power Grid in terms of standard deviation (₹19.48) earning a 1.85% return. Nonetheless, all the three PSUs could not demonstrate a risk-adjusted performance to be attractive on a standalone level in the given time span since Coal India had a modest figure of 0.19.

In comparison to its competitors, however, Coal India has earned its positive Sharpe return alongside the highest absolute volatility in share price. Thus, the performance of the three PSUs demonstrated that they are operating more appropriately in the capacity of mode commodity-linked PSUs involving higher returns mixed with higher volatility (Bhowmik & Wang, 2020). Nonetheless, negative performance of Power Grid indicated that low volatility does not necessarily constitute good Sharpe return since despite being volatile, it resulted in negative return when measured in Sharpe terms. Coal India's 0.19, unremarkable in isolation, was still the only positive risk-adjusted return among the three, earned alongside the group's highest absolute volatility, a combination consistent with commodity-linked PSUs behaving more like leveraged bets on the broader cycle than as defensive holdings (Bhowmik & Wang, 2020). Power Grid's negative Sharpe ratio is a reminder that low volatility and good risk-adjusted return are not the same thing: an asset can be calm and still lose money on a risk-adjusted basis, which is precisely what happened here.

TABLE 1: DESCRIPTIVE, RETURN AND RISK SUMMARY

Stock	Mean (₹)	SD (₹)	Min (₹)	Max (₹)	Total Return (%)	Ann. Volatility (%)	Sharpe Ratio
COALINDIA	416.91	32.21	370.05	481.65	+11.01	22.80	0.19
POWERGRID	286.23	15.30	251.35	320.90	-2.79	20.76	-0.45
ONGC	254.90	19.48	231.33	301.40	+1.85	23.31	-0.20

N = 246 trading days per stock (21 Jul 2025 – 20 Jul 2026); Sharpe ratio assumes $r_f = 6.5\%$ p.a.

The shape of the distribution, however, was different for each series. Coal India's closing prices were relatively flat (platykurtic) with a slight right skew (skewness = 0.20, excess kurtosis = -1.38), which fit nicely into the two trading ranges seen in Figure 1(a): a pre-rally and post-rally trading range. Power Grid's distribution was nearly symmetrical with a moderate peak, given a skewness value of -0.05 and excess

kurtosis of -0.14, consistent with its tighter band trading range. Finally, ONGC's distribution was the most skewed of the three. ONGC had a skewness value of 0.76 and thus its stock prices remained primarily within the ₹230-250 range and were somewhat thin on the right side as it approached the ₹300 range. Statistically, this represents a right-skewed distribution.

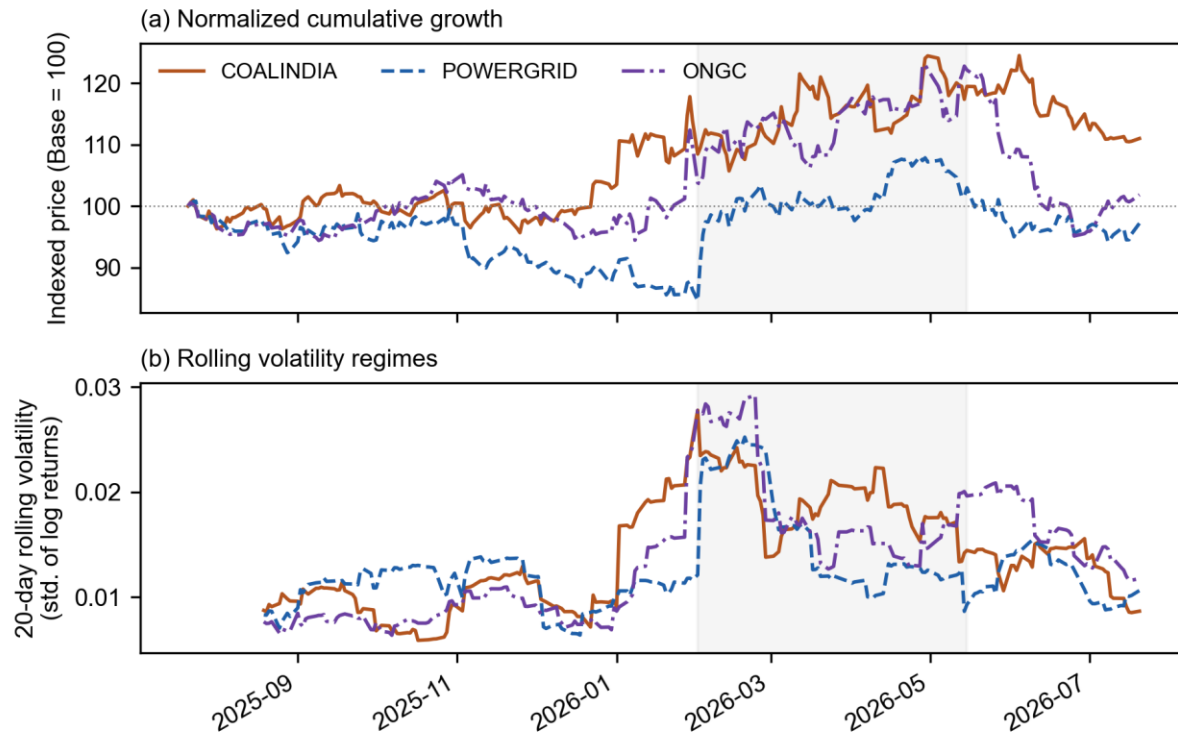
4.2 Co-Movement and Regime Dynamics

Figure 1 (in the following page) shows the normalized growth (base = 100) and volatility 20-day rolling for all three stocks together. Two regimes are clearly visible. From approximately February until around mid-May 2026 (shaded area), all three series were up, and spikes in volatility in the three-week period for all three stocks occurred (Coal India peaked at 0.0278 at 1 February; Power Grid at 0.0252 at 19 February; ONGC at 0.0292 at 23 February), which supports one shared sector-wide shock instead of three independent stories. As for the pairwise correlations throughout the entire period, these were low (Coal India-Power Grid: 0.32; Coal India-ONGC: 0.35; Power Grid-ONGC: 0.25), increased slightly during the time of the rally (0.36/0.28/0.13), and then dropped sharply once the rally ended (0.07/0.17/0.16). In other words, when it comes to their movements in risk (volatility), the three PSUs moved together far more than their movements in direction (returns), which a simple correlation table would not tell.

4.3 Stationarity and Volatility Clustering

In the post-May period, the three stories begin to diverge in the strongest way. ONGC lost nearly all of its gains from its maximum on 29 April when it reached ₹301.40 to minimum of ₹233.10 by June 25, 2026 (22.7% decline) and then partially rebounded toward support at around ₹240-250 by late July. Power Grid and Coal India also pulled back from their own maximums achieved in late April by 11.2% and by 12.8%, however not nearly as much, finishing this period with index values at 111.0 and 97.2 respectively compared to 101.9 for ONGC. This statement is more precise than simply stating that "both stocks have increased substantially." All three PSUs declined but ONGC's correction was nearly twice as large as the others.

The results from the Augmented Dickey-Fuller test are shown in Table 2. The price levels of the three stocks will fit a unit root process ($p = 0.58, 0.30, 0.39$), but the null hypothesis of a unit root can be rejected at a high level of confidence for the log returns ($p < 0.001$). Therefore, the returns are modeled instead of the price levels in Section 3.4, as visually suggested by the growth-index graph.

FIGURE 1: NORMALIZED GROWTH AND VOLATILITY REGIMES

Source: Computed by Author using Python

4.3 Stationarity and Volatility Clustering

In the post-May period, the three stories begin to diverge in the strongest way. ONGC lost nearly all of its gains from its maximum on 29 April when it reached ₹301.40 to minimum of ₹233.10 by June 25, 2026 (22.7% decline) and then partially rebounded toward support at around ₹240-250 by late July. Power Grid and Coal India also pulled back from their own maximums achieved in late April by 11.2% and by 12.8%, however not nearly as much, finishing this period with index values at 111.0 and 97.2 respectively compared to 101.9 for ONGC. This statement is more precise than simply stating that “both stocks have increased substantially.” All three PSUs declined but ONGC’s correction was nearly twice as large as the others. The results from the Augmented Dickey-Fuller test are shown in Table 2. The price levels of the three stocks will fit a unit root process ($p = 0.58, 0.30, 0.39$), but the null hypothesis of a unit root can be rejected at a high level of confidence for the log returns ($p < 0.001$). Therefore, the returns are modeled instead of the price levels in Section 3.4, as visually suggested by the growth-index graph.

TABLE 2: STATIONARITY (ADF) AND VOLATILITY-CLUSTERING (ARCH-LM) DIAGNOSTICS

Stock	ADF Level (p)	ADF Return (p)	Stationary Series	ARCH-LM (p)	Vol. Clustering
COALINDIA	0.584	<0.001	Returns only	0.504	No
POWERGRID	0.303	<0.001	Returns only	0.002	Yes
ONGC	0.389	<0.001	Returns only	0.006	Yes

According to the results of the ARCH-LM test, volatility clustering does not occur in the returns of Coal India ($p = 0.50$) but is present in both Power Grid ($p = 0.002$) and ONGC ($p = 0.006$); thus, in these two time series, one can expect the occurrence of a large number of returns to be followed by another large return in the same direction (Song et al., 2023), while a simple mean return forecast does not predict anything concerning the structure of returns. The test, however, does not address the question of how long this volatility clustering does last in the case of Power Grid and ONGC.

4.4 Forecast Evaluation

Here, we will divert from the usual narrative in the applied ML literature regarding the conclusions of this paper. In Table 3, we will show you the results of cross-validated out-of-sample analysis of the naive random walk benchmark model with the fitted alternatives (i.e., ridge regression; we will omit OLS and shallowly scaled random forests here because those performed no better than the naive benchmark). For each stock examined here, the naive benchmark model's R^2 was less negative (or better) than the fitted alternative's model. Additionally, ahri errors (in rupees) were virtually identical between both approaches. (e.g., Coal India: naive MAE = 5.17; ridge MAE = 5.26; Power Grid: 2.88 vs. 2.99; ONGC 3.24 vs. 3.26).

Moreover, the fitted ridge coefficients can help us understand why. For these three stocks here, no coefficient for any of the eight independent variables (five lags of returns, the rolling volatility, and the RSI) was higher than 0.007 in absolute value; thus, it follows that this model was nearly flat with respect to the final estimated value of each feature. In conclusion, the fitted ridge coefficients provide an explanation of why. An estimation that gives a coefficient of zero is not failing to give a unique signal. Instead, what it is showing is there is no unique linear signal to detect with respect to the independent variables at this time frame. This is quite a different conclusion compared with a poorly configured or improperly specified estimation model, and that is why the results in Table 3 are interpreted in terms of its implications about the marketplace.

TABLE 3: FORECAST ACCURACY – NAIVE BASELINE VS. RIDGE RETURN-MODEL

Stock	Naive R^2	Ridge R^2	Naive MAE (₹)	Ridge MAE (₹)	Next-Day Forecast (₹)	Last Close (₹)
COALINDIA	-0.033	-0.073	5.17	5.26	430.29	429.40
POWERGRID	-0.041	-0.078	2.88	2.99	288.60	288.70
ONGC	-0.015	-0.043	3.24	3.26	249.82	249.58

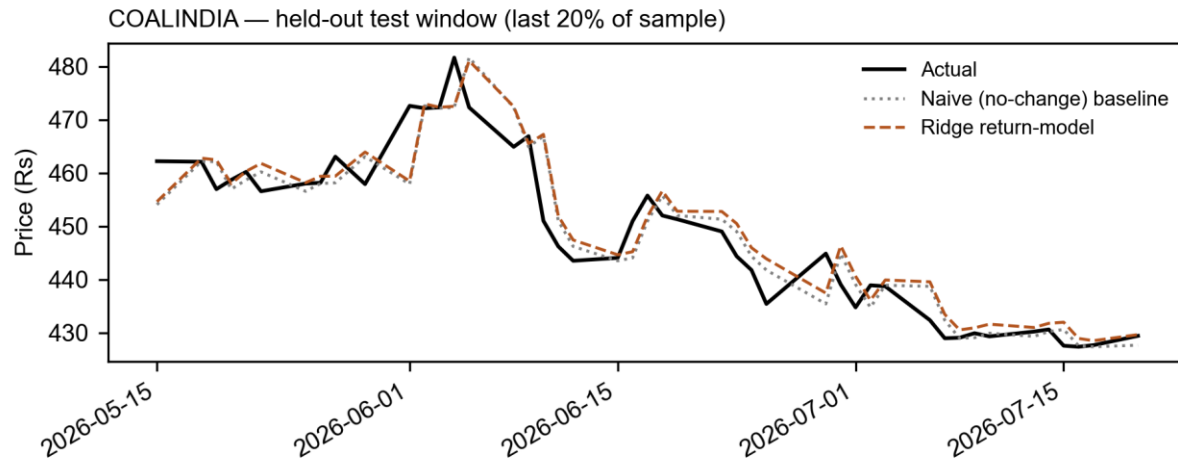
R^2 and MAE are 5-fold expanding-window time-series cross-validation means, evaluated on log returns and converted to price level.

This is also evident in Figure 2 for Coal India: the ridge model forecast line follows the naive line closely when the actual price is tracked for the price of ₹481.65 of Coal India at the beginning of June. However, the ridge method overshoots, thus leading to a larger error compared to using the naive approach. The results of one-step-ahead forecasts for July 21, 2026, reinforce this: ₹430.29 (ridge) is the forecast, while the actual figure is ₹429.40; for Power Grid, the respective figures are ₹288.60 and ₹288.70, and for ONGC, it is ₹249.82 against ₹249.58. Since the differences between the forecast and the actual price is negligible, they do not affect practical decision-making in any way.

Nonetheless, all of the above does not mean the feature engineering process did not yield positive returns. Daily log returns for big liquid PSU stocks behave in an unpredictable manner as regards their conditional mean, which is what the weak form of market efficiency suggests (Fama, 1970). If a predictor beats this naive benchmark significantly through a naïve benchmark, then it is usually an indication of data leakage and nothing else – which is exactly what will be produced through a shuffled non-time-ordered train-test split. Rather, the lagged prices and technical indicators do not contain exploitable information about the future prices of these three PSUs in question, however, the ARCH-LM evidence in Section 3.3 signifies that the aforementioned variables have information regarding risk. Therefore, the actionable information that can be derived from this is effective control of risk that should be utilized rather than making predictions about price movements.

4.5 Volatility Persistence: GARCH(1,1) Estimates

Table 4 shows the GARCH(1,1) fitting for each of the stocks. The ARCH coefficient of Coal India is statistically not different from zero ($\alpha = 0.035$, $p = 0.153$), although its GARCH coefficient is large and accurately estimated, that is $\beta = 0.945$, $p < 0.001$. Because a very high and statistically significant β can occur on a process which displays no real clustering of shocks as long as it has been fitted with a very slow-moving variance trend, this result serves to reinforce the findings in Section 3.3 rather than weakens them.

FIGURE 2: ACTUAL VS. NAIVE BASELINE VS. RIDGE RETURN-MODEL (HELD-OUT TEST WINDOW)

Source: Computed by Author using Python

TABLE 4: GARCH(1,1) PARAMETER ESTIMATES AND IMPLIED VOLATILITY HALF-LIFE

Stock	α (ARCH)	β (GARCH)	Persistence ($\alpha+\beta$)	Half-Life (days)	Ann. Unconditional Vol. (%)
COALINDIA	0.035 (p=.153)	0.945 (p<.001)	0.980	33.8	23.5
POWERGRID	0.354 (p=.092)	0.020 (p=.919)	0.374	<1	21.1
ONGC	0.052 (p=.167)	0.943 (p<.001)	0.996	158.0	36.1

Estimated by maximum likelihood, constant-mean specification, normal innovations. *p*-values in parentheses; half-life = $\ln(0.5)/\ln(\text{persistence})$.

The ARCH-LM test in Section 3.3 has shown that both Power Grid and ONGC have surpassed this hurdle, but their GARCH fits indicate that they achieved this feat in almost completely opposite fashion. The ARCH term for Power Grid is the highest among the three ($\alpha = 0.354$). However, the GARCH term for this stock is almost zero, which means that shocks arrive with a high intensity but then fade away quickly, resulting in a half-life of less than one trading day. In contrast, ONGC has a low and insignificant ARCH term ($\alpha = 0.052$, $p = 0.167$). However, the GARCH term is high and significant ($\beta = 0.943$, $p < 0.001$), implying that the total persistence comes to 0.996, which is close to the unit root threshold of 1.0 and leads to a halving of the intensity of the volatility shock after some 158 trading days, which is almost a year. Figure 3 (in the following page) illustrates this difference in a clearer manner than what can be seen from the 20-day rolling window in Figure 1(b). Power Grid has reached a level of high conditional volatility (indicated with blue) that fades away quickly, while Coal India (orange) and ONGC (purple) display a slow decrease in the level of volatility, which has been seen in detail in Section 3.2.

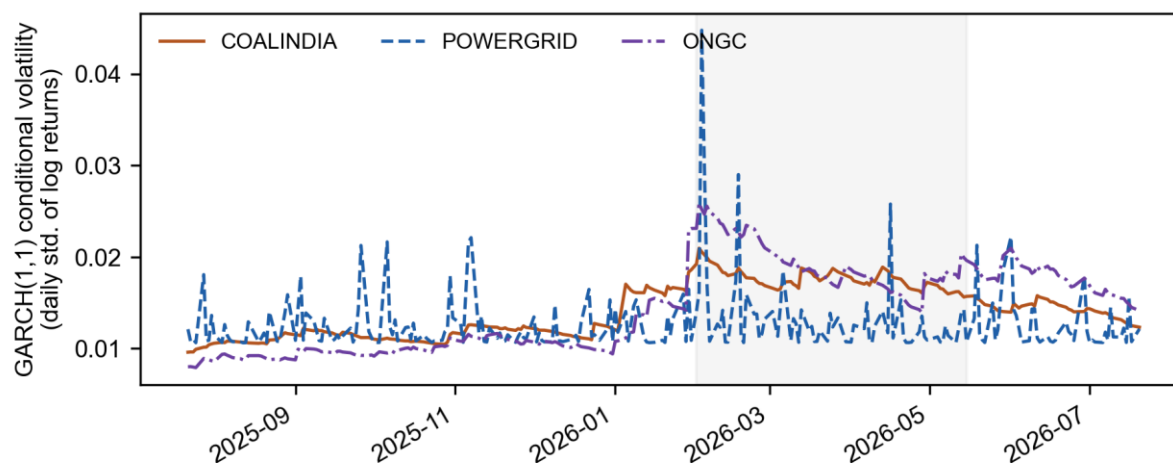
5. IMPLICATIONS AND LIMITATIONS

5.1 Implications

The clearest implication is that these three PSUs, despite common state ownership, do not behave as one homogeneous risk category. Coal India offers the highest return but also the highest and least clustered volatility; Power Grid offers low volatility alongside a negative Sharpe ratio over this window; and ONGC combines moderate average volatility with the deepest single drawdown and the only clearly right-skewed price distribution of the three. Treating “energy PSU” as a single exposure would mask all of this.

There is also a methodological lesson: benchmark against a naive baseline under a time-respecting split before reporting any accuracy gain, since a market efficient in the mean but not the variance, exactly what the combination of a random-walk-beating forecast and clustering in two of three stocks points to here, argues for GARCH-type models over point forecasting. Section 3.5's estimates go further, showing Power Grid and ONGC clear the same clustering test for almost opposite reasons, one fast-decaying, one

FIGURE 3: GARCH(1,1) CONDITIONAL VOLATILITY BY STOCK



Source: Computed by Author using Python

near-permanent. For practitioners, position sizing and stop levels are better supported by this data than directional trading signals are.

5.2. Limitations

The analysis covers a single one-year window and three tickers within one sub-sector, which limits how far the conclusions generalise to other periods, horizons, or industries. The forecasting exercise is restricted to a one-day-ahead horizon using price-derived features only, without macroeconomic, news, or order-flow variables that might carry information the technical feature set does not. The Sharpe ratio calculation also rests on a single assumed risk-free rate rather than a matched-maturity series. The GARCH(1,1) models assume symmetric normal innovations and cannot capture the leverage effect through which equal-sized positive and negative shocks move volatility differently; Power Grid's near-zero, insignificant GARCH term also means its half-life estimate should be read as indicative rather than precise.

Future Scope

Extending the sample across multiple years and market cycles would clarify whether the efficiency result documented here is stable or specific to the window studied. An asymmetric specification such as EGARCH or GJR-GARCH would test whether ONGC's and Coal India's volatility responds differently to negative shocks than to positive ones, a question this symmetric GARCH(1,1) cannot answer, alongside multi-day-horizon forecasts and sentiment or macro-news features to test whether they add anything beyond the price-derived indicators used here.

6. SUMMARY AND CONCLUSION

Coal India, Power Grid, and ONGC turn out to be a useful natural experiment in how much “same sector, same ownership structure” actually predicts about price behaviour: not very much, once return, volatility, correlation, and distributional profiles are compared side by side. The stationarity results and the out-of-sample forecasting exercise point the same way, toward genuine unpredictability in price direction at a one-day horizon rather than a modelling shortfall. What is left over is the volatility-clustering signal in two of the

three stocks, and the GARCH(1,1) estimates show that signal is not one thing but two: a fast-decaying risk state in Power Grid and a near-permanent one in ONGC.

REFERENCES

- Bhowmik, R. and Wang, S., "Stock Market Volatility and Return Analysis: A Systematic Literature Review," *Entropy*, Volume 22, Number 5, Page 522, 2020.
- Chatterjee, A., Bhowmick, H., and Sen, J., "Stock Price Prediction Using Time Series, Econometric, Machine Learning, and Deep Learning Models," 2021 IEEE Mysore Sub Section International Conference (MysuruCon), Pages 289-296, 2021.
- Chauhan, A., Mayur, P., Gokarakonda, Y. S., Jamie, P., and Mehrotra, N., "Indian Stock Market Prediction using Augmented Financial Intelligence ML," SSRN Working Paper, 2024.
- Dickey, D. A. and Fuller, W. A., "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, Volume 74, Number 366, Pages 427-431, 1979.
- Engle, R. F., "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, Volume 50, Number 4, Pages 987-1007, 1982.
- Fama, E. F., "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, Volume 25, Number 2, Pages 383-417, 1970.
- He, K., Yang, Q., Ji, L., Pan, J., and Zou, Y., "Financial Time Series Forecasting with the Deep Learning Ensemble Model," *Mathematics*, Volume 11, Number 4, Page 1054, 2023.
- Kim, Y., Yoo, S., and Park, S., "A Rule-Based Stock Trading Recommendation System Using Sentiment Analysis and Technical Indicators," *Electronics*, Volume 14, Number 4, Page 773, 2025.
- Lee, M.-C., Chang, J.-W., Yeh, S.-C., Chia, T.-L., Liao, J.-S., and Chen, X.-M., "Applying Attention-Based BiLSTM and Technical Indicators in the Design and Performance Analysis of Stock Trading Strategies," *Neural Computing and Applications*, Volume 34, Pages 13267-13279, 2022.
- Li, Y. and Pan, Y., "A Novel Ensemble Deep Learning Model for Stock Prediction Based on Stock Prices and News," *International Journal of Data Science and Analytics*, Volume 13, Pages 139-149, 2021.
- Liang, L., "ARIMA with Attention-Based CNN-LSTM and XGBoost Hybrid Model for Stock Prediction in the US Stock Market," *SHS Web of Conferences*, Volume 196, Page 02001, 2024.
- Nabipour, M., Nayyeri, P., Jabani, H., and Mosavi, A., "Deep Learning for Stock Market Prediction," *Entropy*, Volume 22, Number 8, Page 840, 2020.
- Rammurthy, S. K. and Patil, S. B., "An LSTM-Based Approach to Predict Stock Price Movement for IT Sector Companies," *International Journal of Cognitive Informatics and Natural Intelligence*, Volume 15, Number 4, Pages 1-12, 2021.
- Sang, S. and Li, L., "A Novel Variant of LSTM Stock Prediction Method Incorporating Attention Mechanism," *Mathematics*, Volume 12, Number 7, Page 945, 2024.
- Sen, J., Dutta, A., and Mehtab, S., "Stock Portfolio Optimization Using a Deep Learning LSTM Model," 2021 IEEE Mysore Sub Section International Conference (MysuruCon), Pages 263-271, 2021.

- Sheppard, K., "arch: Python Module for Conditional Heteroskedasticity Models," Version 8.0.0, Zenodo, <https://doi.org/10.5281/zenodo.605835>, 2025.
- Singh, T., Kalra, R., Mishra, S., Satakshi, and Kumar, M., "An Efficient Real-Time Stock Prediction Exploiting Incremental Learning and Deep Learning," *Evolving Systems*, Volume 14, Pages 919-937, 2022.
- Song, Y., Tang, X., Wang, H., and Ma, Z., "Volatility Forecasting for Stock Market Incorporating Macroeconomic Variables Based on GARCH-MIDAS and Deep Learning Models," *Journal of Forecasting*, Volume 42, Number 1, Pages 51-59, 2023.
- Wang, J. and Kim, J., "Predicting Stock Price Trend Using MACD Optimized by Historical Volatility," *Mathematical Problems in Engineering*, Volume 2018, Pages 1-12, 2018.

REALIZED KURTOSIS AND VOLATILITY FORECASTING BEYOND VIX-IMPLIED VARIANCE

Megan Y. Sun, School of Business and Economics, University of Wisconsin – River Falls, U.S.A.
Dawn Hukai, School of Business and Economics, University of Wisconsin – River Falls, U.S.A.

dx.doi.org/10.18374/JABE-26-3.6

ABSTRACT

Using E-mini S&P 500 futures data from 2003–2026, we find that realized kurtosis contains persistent predictive information for future realized volatility. The predictive power survives controls for autoregressive volatility dynamics, lagged returns, and option-implied volatility. While implied variance captures forward-looking market expectations, realized kurtosis contributes additional information related to persistent tail-risk dynamics. However, realized skewness and the Amihud illiquidity measure lose predictive power for realized volatility when realized kurtosis is present. We find evidence that a HAR model augmented with multiple lagged returns, implied variance, and multiple lagged realized kurtosis provides the most effective framework for forecasting realized volatility. This result indicates that tail-risk persistence represents an important and previously underappreciated determinant of future volatility.

Keywords: HAR-RV model, realized volatility, realized kurtosis, VIX

1. INTRODUCTION

Volatility is one of the most fundamental concepts in finance. It plays a central role in asset pricing, portfolio allocation, risk management, derivative valuation, and financial regulation. Investors, portfolio managers, and policymakers all rely on accurate volatility estimates to assess market risk and make informed decisions. Consequently, understanding the dynamics of volatility and improving volatility forecasts have long been important objectives in financial economics. The increasing availability of high-frequency financial data has popularized the concept of realized volatility, which aggregates information from high-frequency returns and provides a more accurate estimate of the actual variation in asset prices over a given period. Since the seminal work of Andersen, Bollerslev, Diebold, and Labys (2003), realized volatility has become one of the most widely used measures of *ex post* market uncertainty and has generated a large literature devoted to understanding its statistical properties and forecasting performance.

Accurate forecasts of realized volatility are particularly important because volatility exhibits strong persistence and clustering. Periods of elevated volatility are often followed by additional periods of heightened uncertainty, while tranquil market conditions tend to persist over time. This persistence creates opportunities for predictive modeling and has motivated the development of numerous volatility forecasting frameworks, including GARCH-type models, stochastic volatility models, and the heterogeneous autoregressive (HAR) model introduced by Corsi (2009). Among these approaches, the HAR model has become especially popular due to its simplicity, economic intuition, and ability to capture long-memory features observed in realized volatility.

Despite the success of existing forecasting models, an important question remains as to whether additional information beyond past realized volatility can improve forecasts of future volatility. In particular, measures derived from higher-order return moments and option markets may contain valuable information about future market uncertainty. Understanding which variables provide incremental predictive content beyond realized volatility is important both from a practical forecasting perspective and from a broader economic perspective, as it sheds light on the mechanisms through which information about risk is transmitted across financial markets.

This study investigates whether option-implied variance and realized kurtosis contain predictive information for future realized volatility beyond that captured by the HAR model. Our research contributes to empirical finance in the following ways. First, this study is among the first to investigate E-mini S&P 500 futures (ES),

which has nearly continuous trading, high liquidity, and lower microstructure noise. The data also spans a very long time period from the introduction of VIX in 2003 to April 2026, a period covering the late dot-com era, the Great Financial Crisis of 2007-2009, flash crashes, Covid, and recent inflation shocks. In addition, high frequency data is used to construct 5-minute realized measures instead of daily data.

Second, this study is the first to investigate if realized kurtosis contains predictive information for future realized volatility beyond that already embedded in option-implied volatility. While implied variance derived from VIX remains the single most important predictor in the model, realized kurtosis continues to exhibit significant explanatory power after controlling for option-implied expectation for volatility. This result suggests that realized kurtosis captures aspects of tail-risk dynamics that are not fully reflected in option prices.

Furthermore, the superior performance of the identified augmented HAR specification indicates that tail-risk information persists across several trading days. Rather than affecting volatility only contemporaneously, extreme return movements appear to influence volatility expectations over a longer horizon with a decaying power. These findings highlight the importance of incorporating higher-order realized moments into volatility forecasting models and suggest that persistent tail-risk measures can enhance both in-sample model fit and out-of-sample forecast accuracy.

This paper is organized as follows. Section 2 reviews previous research. In Section 3, the data and methodology used in this study are described. Section 4 presents empirical results. Section 5 concludes with a summary of evidence.

2. LITERATURE REVIEW

A growing body of research suggests that higher-order realized moments contain information about market tail risk and jump activity that is not fully captured by realized volatility alone. Because realized kurtosis assigns greater weight to extreme return realizations, it provides a natural measure of tail risk and jump activity and thus may contain predictive information about future volatility beyond that embedded in historical volatility measures or option-implied expectations. This study contributes to the literature by examining whether realized kurtosis retains incremental forecasting power after controlling for implied variance within a heterogeneous autoregressive (HAR) framework.

The modern realized volatility literature originates with Andersen, Bollerslev, Diebold, and Labys (2003), who demonstrate how high-frequency intraday returns can be aggregated to construct accurate measures of daily realized volatility. Using foreign exchange market data, they show that realized volatility provides a substantially more precise estimate of latent volatility than measures based solely on daily returns. Furthermore, they document that the logarithm of realized volatility exhibits strong persistence and approximates Gaussian behavior, making it well suited for econometric modeling and forecasting. Their work established realized volatility as a practical measure of *ex post* market uncertainty and laid the foundation for a large subsequent literature on volatility forecasting.

Subsequent studies emphasized the importance of jumps and extreme price movements in explaining volatility dynamics. Barndorff-Nielsen and Shephard (2004) develop the concepts of power variation and bipower variation, providing a framework for decomposing realized volatility into continuous and jump components. Their results demonstrate that discontinuous price movements contribute significantly to total return variation and suggest that measures sensitive to jumps may contain valuable information about future volatility. Similarly, Huang and Tauchen (2005) document that jump variation accounts for a substantial portion of realized volatility, particularly during periods of market stress. Together, these studies highlight the importance of tail events and jump activity and provide a strong economic motivation for incorporating higher-order realized moments into volatility forecasting models.

Building on the realized volatility literature, Corsi (2009) proposes the Heterogeneous Autoregressive (HAR) model, which approximates the long-memory behavior of realized volatility through daily, weekly, and monthly volatility components. Due to its simplicity, economic intuition, and strong forecasting

performance, the HAR model has become one of the most widely used frameworks for modeling and forecasting realized volatility. Consequently, the HAR model serves as the benchmark specification in the present study.

A parallel literature investigates the measurement and economic significance of higher-order realized moments. Neuberger (2012) develops methods for extracting realized skewness directly from high-frequency returns and demonstrates that realized skewness is both time-varying and predictable. Extending this line of research, Kozhan, Neuberger, and Schneider (2013) introduce skew swaps and document the existence of a skewness risk premium, providing evidence that investors require compensation for bearing downside tail risk. Their findings suggest that higher-order moments contain economically meaningful information beyond that reflected in variance alone.

Additional evidence on the economic relevance of realized moments is provided by Amaya, Christoffersen, Jacobs, and Vasquez (2015), who use intraday data to construct weekly realized moments and show that realized skewness and realized kurtosis are significant predictors of future stock returns. Their results indicate that higher-order moments capture important information about market conditions that is not contained in realized volatility. More recently, Bae and Lee (2021) propose new realized third- and fourth-order comoments that incorporate both volatility-of-volatility effects and jump contributions. They show that these higher-order measures capture important features of financial markets, including volatility persistence and return-volatility asymmetry. Neuberger and Payne (2021) further demonstrate that long-horizon skewness and kurtosis can be estimated accurately using short-horizon returns and leverage effects, providing additional support for the use of high-frequency data in constructing realized higher moments.

While the studies above establish the importance of realized higher moments in predicting returns, relatively few papers examine their role in volatility forecasting. One notable exception is Mei, Liu, Ma, and Chen (2017), who incorporate realized skewness and realized kurtosis into variants of the HAR model and show that higher-order moments can improve volatility forecasts. Their results provide direct evidence that realized moments contain information beyond realized volatility itself. However, their analysis does not investigate whether realized kurtosis retains predictive power after controlling for information embedded in option markets, nor does it explore the persistence of realized kurtosis through alternative lag structures.

Bonato, Cepni, Gupta, and Pierdzioch (2021) extend this line of research by investigating the role of realized skewness and realized kurtosis in forecasting realized volatility across international real estate investment trust (REIT) markets. Using variants of the HAR model, they find that higher-order realized moments improve volatility forecasts in several international markets, although the magnitude of the forecasting gains differs across countries. Their results provide additional evidence that realized higher moments contain economically meaningful information for forecasting future volatility.

Recent research by Liang, Wei, and Zhang (2020) shifts the focus toward the role of option-implied information in volatility forecasting. Using realized volatility and implied volatility from eight major international stock markets within a HAR framework, they show that implied-volatility-based measures significantly improve volatility forecasts across multiple prediction horizons. While their findings establish implied volatility as an important benchmark for forecasting realized volatility, they do not examine whether realized higher moments continue to provide incremental forecasting information once implied volatility has been incorporated into the forecasting model. The present study addresses this question by investigating whether realized skewness and kurtosis retains predictive power beyond VIX-implied variance within a HAR framework.

This study extends the existing literature in several important ways. First, it examines whether realized kurtosis contributes incremental forecasting power beyond implied variance, a forward-looking measure of market expectations derived from VIX. Second, it investigates the persistence of tail-risk information by considering both multiple lag specifications and aggregated measures of realized kurtosis. Finally, the analysis evaluates the contribution of realized kurtosis using both in-sample and out-of-sample forecasting performance within a HAR framework. The empirical results indicate that realized kurtosis remains a significant predictor of future realized volatility even after controlling for option-implied information, and that its predictive effects persist across multiple trading days.

3. DATA AND METHODOLOGY

3.1 Data and Variable Construction

This study examines the relationship between realized kurtosis, implied variance, and future realized volatility using S&P 500 E-mini (ES) futures and the CBOE Volatility Index (VIX). Although ES futures data are available beginning in 1997, the sample period in this study is restricted to September 22, 2003 through April 30, 2026. The starting date coincides with the introduction of the modern CBOE VIX methodology. Intraday ES futures and VIX data are obtained from Pi Trading. Because ES futures trade nearly twenty-four hours per day, the choice of trading session can influence realized volatility estimates. Overnight trading tends to exhibit lower liquidity and has experienced substantial structural changes over the sample period. To maintain consistency and reduce the influence of evolving overnight market conditions, all baseline realized measures are constructed using regular trading hours Monday through Friday 9:30 a.m. to 4:00 p.m. Eastern time.

The use of high-frequency data introduces potential market microstructure biases, including bid-ask bounce and nonsynchronous trading effects. Andersen et al. (2003) demonstrate that volatility estimates based on very high-frequency returns may be distorted by such effects and recommend the use of five-minute returns as a practical compromise between measurement precision and microstructure bias. Accordingly, all realized measures in this study are constructed using five-minute intraday returns. The i^{th} five minute return on ES futures contract is calculated using the following equation:

$$r_{t,i} = \log(P_{t,i}) - \log(P_{t,i-1}) \quad (1)$$

where $r_{t,i}$ is the intraday i^{th} five minutes return of the ES futures contract at i^{th} interval on day t , $P_{t,i}$ is the ending futures price for the i^{th} interval on day t , while $P_{t,i-1}$ is the ending futures price for the $(i-1)^{\text{th}}$ interval on day t . The daily return on ES futures contract is calculated using the following equation:

$$R_t = \log(P_t) - \log(P_{t-1}) \quad (2)$$

where R_t is the return of the ES futures contract on day t , P_t is the closing price of ES futures on day t , while P_{t-1} is the closing price of the ES futures contract on day $t-1$.

Daily realized volatility is obtained by aggregating squared five-minute returns. Following Andersen and Bollerslev (1998), realized volatility provides a nonparametric estimate of the latent integrated volatility process. The realized volatility is thus calculated using the following equation.

$$RV_t = \sum_{i=1}^M r_{t,i}^2 \quad (3)$$

where RV_t is the daily realized volatility of the ES futures on day t . $r_{t,i}$ is the intraday i^{th} five minutes return of the ES futures contract at i^{th} interval on day t , and M is the number of intraday five minutes intervals. Because realized volatility is highly right skewed, the logarithm of realized volatility is used throughout the empirical analysis. Log transformation is standard in the realized volatility literature and produces a distribution that is closer to Gaussian while improving regression stability. For expositional convenience, the term realized volatility (RV) refers to the logarithm of realized volatility unless otherwise noted.

In addition to realized volatility, this study considers several measures designed to capture higher-order return characteristics and market conditions. Realized skewness measures return asymmetry and reflects the tendency for large positive or negative return realizations. Realized kurtosis captures tail risk and the prevalence of extreme return movements. Because realized skewness and kurtosis are sensitive to scale, normalized realized skewness (NRS) and normalized realized kurtosis (NRK) are employed following the existing literature. These normalized measures facilitate comparisons across time and market conditions. The realized skewness and the realized kurtosis are calculated using the following equation.

$$RS_t = \sum_{i=1}^M r_{t,i}^3 \quad (4)$$

$$RK_t = \sum_{i=1}^M r_{t,i}^4 \quad (5)$$

where RS_t is the realized skewness, RK_t is the realized kurtosis, $r_{t,i}$ are the intraday i^{th} five minutes returns for day t , and M is the number of intraday five minutes intervals. A normalized realized skewness and kurtosis often used in finance literature are

$$NRS_t = \frac{\sqrt{M} \cdot RS_t}{RV_t^{3/2}} \quad (6)$$

$$NRK_t = \frac{M \cdot RK_t}{RV_t^2} \quad (7)$$

where M is the number of intraday five minutes intervals, and RV is the realized volatility without logarithmic transformation we calculated from equation (3).

Market liquidity is measured using the Amihud (2002) illiquidity measure. The Amihud (2002) illiquidity measure is one of the most widely used liquidity proxies in the finance literature. It captures the price impact associated with trading activity and is commonly interpreted as a proxy for market depth and trading frictions. Higher values indicate that relatively small trading volumes are associated with larger price movements, suggesting reduced market liquidity. To improve numerical stability and interpretability, the Amihud measure is scaled and subsequently transformed using logarithms in the empirical analysis. Amihud illiquidity is calculated as following.

$$\text{Amihud}_t = \frac{1}{M} \sum_{i=1}^M \frac{|r_{t,i}|}{\text{Closing Price} \cdot 50 \cdot \text{Volume}_{t,i}} \quad (8)$$

where Volume is number of contracts traded in each 5-minute interval. The Amihud illiquidity measure is scaled by e^{10} to make the number easier to read.

Forward-looking volatility expectations are captured using the VIX index. Since the VIX represents annualized implied volatility, it is converted into an implied variance measure to facilitate direct comparison with daily realized volatility. Specifically, implied variance (IV) is constructed from the VIX level and subsequently transformed using logarithms. This transformation aligns the scale of implied variance with the dependent variable used in the forecasting regression models. Since the VIX is quoted as an annualized volatility:

$$VIX = 100 \times \sqrt{252 \times IV_t} \quad (9)$$

where IV_t is the expected daily implied variance under the risk-neutral measure. Rearranging, we have the IV calculated using the following equation.

$$IV_t = \left(\frac{VIX_t}{100} \right)^2 \frac{1}{252} \quad (10)$$

In addition to implied variance, this study considers the realized volatility of the VIX itself (VIXRV). VIXRV is computed by aggregating squared intraday VIX returns and is intended to capture volatility-of-volatility effects. Similar to the realized volatility and implied variance derived from VIX, the logarithm of VIXRV is employed in the empirical analysis to reduce skewness and improve statistical properties. VIXRV is calculated using the following equation.

$$vixr_{t,i} = \log(VIX_t) - \log(VIX_{t-1}) \quad (11)$$

$$VIXRV_t = \sum_{i=1}^{M_t} vixr_{t,i}^2 \quad (12)$$

3.2 Model Specifications

To forecast future realized volatility, this study employs the Heterogeneous Autoregressive model of realized volatility (HAR-RV) proposed by Corsi (2009). The HAR framework is designed to approximate the long-memory behavior commonly observed in financial volatility through the interaction of daily, weekly, and monthly volatility components. Due to its simplicity, economic intuition, and strong forecasting performance, the HAR model has become one of the most widely used approaches in the realized volatility literature and serves as the benchmark specification in this study.

First, following Corsi (2009), the baseline HAR model is constructed to relates future realized volatility to lagged daily, weekly, and monthly realized volatility components.

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \epsilon_{t+1} \quad (13)$$

where RV_t is daily realized volatility on day t , RV_t^w is last week's average volatility, and RV_t^m is the last month's average volatility. Specifically,

$$RV_t^w = \frac{1}{5} \sum_{i=1}^5 RV_{t+1-i} \quad (14)$$

$$RV_t^m = \frac{1}{22} \sum_{i=1}^{22} RV_{t+1-i} \quad (15)$$

This study first augments the baseline HAR model by incorporating lagged daily returns to account for potential leverage and return-volatility effects.

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{i=1}^5 \beta_i R_{t+1-i} + \epsilon_{t+1} \quad (16)$$

The analysis then evaluates whether higher-order realized moments, including normalized realized skewness and normalized realized kurtosis, provide incremental explanatory power beyond the persistence captured by the HAR structure. The forecasting framework is also further extended by incorporating the Amihud illiquidity measure to assess whether market liquidity conditions contribute to volatility dynamics.

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{i=1}^5 \beta_i R_{t+1-i} + \theta_s NRS_t + \epsilon_{t+1} \quad (17)$$

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{i=1}^5 \beta_i R_{t+1-i} + \theta_k NRK_t + \epsilon_{t+1} \quad (18)$$

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{i=1}^5 \beta_i R_{t+1-i} + \theta_a Amihud_t + \epsilon_{t+1} \quad (19)$$

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{i=1}^5 \beta_i R_{t+1-i} + \theta_s NRS_t + \theta_k NRK_t + \theta_a Amihud_t + \epsilon_{t+1} \quad (20)$$

Subsequently, implied variance derived from the VIX is added to determine whether forward-looking information from option markets improves volatility forecasts. Because implied variance reflects market expectations regarding future uncertainty, it provides a natural benchmark against which the incremental forecasting power of realized kurtosis can be evaluated.

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{i=1}^5 \beta_i R_{t+1-i} + \theta_v IV_t + \theta_{rv} VIXRV_t + \epsilon_{t+1} \quad (21)$$

Since the primary objective of this study is to determine whether realized skewness and realized kurtosis contains information about future realized volatility beyond that already embedded in implied variance and historical volatility persistence, models including both implied variance and realized kurtosis are estimated and compared with alternative specifications.

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{i=1}^5 \beta_i R_{t+1-i} + \theta_v IV_t + \theta_k NRK_t + \epsilon_{t+1} \quad (22)$$

Finally, the persistence of tail-risk information is examined using alternative realized kurtosis specifications. In addition to a single-lag realized kurtosis measure, the analysis considers both multiple lagged values of realized kurtosis and a weekly average realized kurtosis measure. These alternative specifications allow us to investigate whether the predictive effects of realized kurtosis are short-lived or persist across multiple trading days.

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{i=1}^5 \beta_i R_{t+1-i} + \theta_v IV_t + \sum_{i=1}^5 \theta_{k_i} NRK_{t+1-i} + \epsilon_{t+1} \quad (23)$$

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{i=1}^5 \beta_i R_{t+1-i} + \theta_v IV_t + \theta_w NRK_t^w + \epsilon_{t+1} \quad (24)$$

where $NRK_t^w = \frac{1}{5} \sum_{i=1}^5 NRK_{t-i}$

4. EMPIRICAL RESULTS

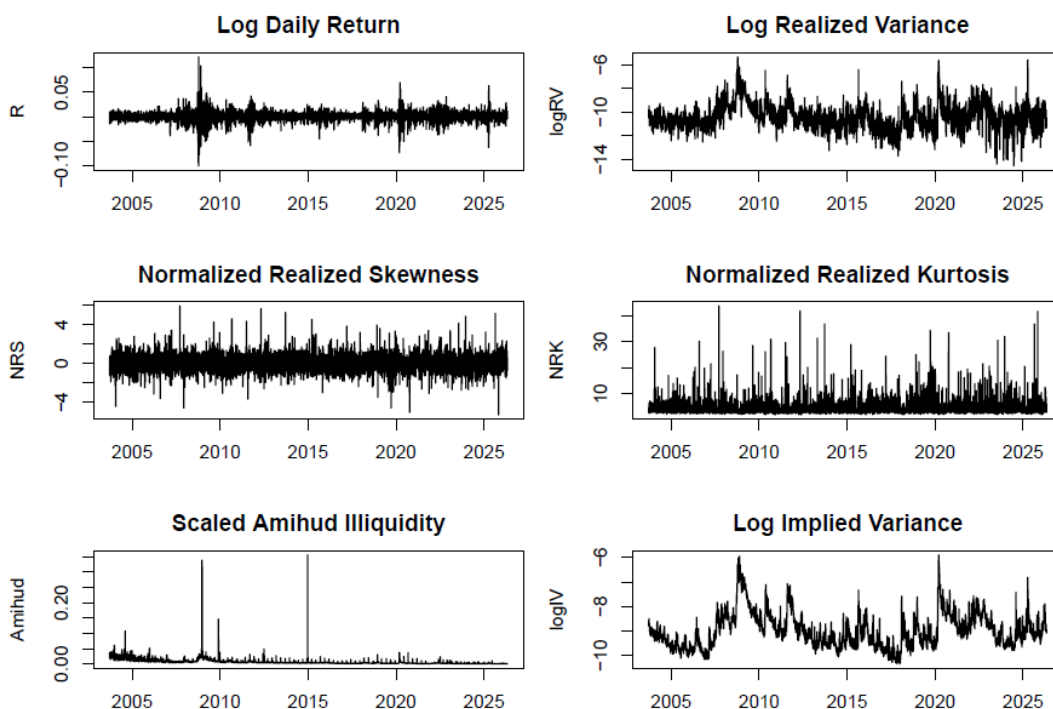
As a preliminary step, this study examines the time-series behavior of the key variables used in the empirical analysis. Figure 1 presents the daily return, realized volatility, normalized realized skewness, normalized realized kurtosis, Amihud illiquidity, and implied variance series over the sample period. Several well-known facts emerge from the data. First, both realized volatility and implied variance exhibit strong persistence and volatility clustering. Periods of elevated volatility are followed by additional periods of heightened uncertainty, particularly during episodes of market stress such as the Global Financial Crisis and the COVID-19 pandemic. The close comovement between realized volatility and implied variance provides preliminary evidence that option-implied information may be useful for forecasting future realized volatility.

Second, normalized realized kurtosis displays substantial time variation and occasional sharp spikes, indicating the presence of extreme return realizations and heightened tail risk. These spikes tend to coincide with periods of market turbulence and suggest that realized kurtosis may contain information about future volatility beyond that captured by realized volatility alone. In contrast, normalized realized skewness fluctuates around zero and exhibits considerably less persistence, which is consistent with previous studies documenting weaker time-series dependence in skewness measures.

Third, the Amihud illiquidity measure remains close to zero for most of the sample period but exhibits several isolated spikes. Examination of the largest observations indicates that many occur during holiday trading sessions and other periods characterized by unusually low trading activity. These findings suggest that extreme Amihud values in the ES futures market are often associated with temporary reductions in trading volume rather than systemic market disruptions. Overall, the relatively low level of illiquidity throughout the sample is consistent with the highly liquid nature of the ES futures market.

Taken together, the graphical evidence suggests that realized volatility, implied variance, and realized kurtosis exhibit the strongest persistence and variation over time. These characteristics make them natural candidates for inclusion in volatility forecasting models and provide preliminary motivation for examining whether realized kurtosis contributes incremental forecasting information beyond implied variance and historical volatility persistence.

Figure 1: Plots of the time series



Before estimating the volatility forecasting models, we also examine the time-series properties of the variables using the Augmented Dickey-Fuller (ADF) test. The null hypothesis of the ADF test is that the series contains a unit root and is therefore nonstationary. Table 1 reports the ADF test statistics which strongly reject the null hypothesis of a unit root for all variables considered. The ADF test statistics range from -5.28 for implied variance to -18.13 for daily returns, with all statistics significant at the 0.1% level. These findings indicate that daily returns, realized volatility, realized higher moments, liquidity measures, and option-implied volatility measures are stationary over the sample period. The stationarity of the variables is important for the subsequent forecasting analysis because it reduces concerns regarding spurious regression results and ensures that the estimated relationships reflect genuine dynamic interactions rather than common stochastic trends. Consequently, all variables are retained in levels for empirical analysis without further differencing or detrending.

Table 1: ADF test for stationarity

This table presents t-statistics for the ADF test. The null hypothesis for the ADF tests is that each series exhibits a unit root. ***, **, and * denote significance at the 0.1%, 1%, 5% level, respectively.

Panel A: S&P 500 Futures	ADF Test
R	-18.1330***
LogRV	-7.4863***
NRS	-16.9334***
NRK	-15.4493***
Amihud	-17.3432***
IV	-5.2796***
IV_RV	-10.5471***

Next, we augment the original HAR model by including lagged daily returns. The regression results are shown in Table 2. Each HAR component including the previous day realized volatility, weekly and monthly

components are all highly significant, even after controlling the lagged daily returns. Note that the realized volatility is positively related to the previous daily, weekly, and monthly realized volatility components, indicating the presence of strong volatility persistence. Also worth noting is the lagged past returns, which are not included in the original HAR model, are negatively related to realized volatility, indicating that a highly negative return will lead to higher future volatility. However, the impact of lagged returns on realized volatility decays over time and essentially disappears after four days. Although the five-lag specification achieves a marginally lower Akaike Information Criterion (AIC), it does have a higher Bayesian Information Criterion (BIC) with almost 6 points (note that BIC does penalize extra parameters). In addition, the fifth lag of daily returns is statistically insignificant. Moreover, the four-lag return specification yields a lower BIC and virtually identical adjusted R^2 . Following the principle of parsimony, we adopt the four-lag specification (HAR+Ret4) as the benchmark model.

Table 2: HAR model with lagged returns

This table reports the regression result of the original HAR model and augmented models with lagged daily returns. Coefficients are reported together with Newey-West standard error in parentheses. ***, **, and * denote significance at the 0.1%, 1% , 5% level, respectively. The parentheses by model indicate the equations specified in section 3.2.

Variable	Model (13) Original HAR	Model (16a) HAR+Ret5	Model (16b) HAR+Ret4
Intercept	-0.6154*** (0.1103)	-0.7309*** (0.0892)	-0.7217*** (0.0890)
HAR_D	0.4212*** (0.0224)	0.2989*** (0.0228)	0.2977*** (0.0230)
HAR_W	0.3847*** (0.0322)	0.4229*** (0.0309)	0.4339*** (0.0302)
HAR_M	0.1348*** (0.0253)	0.2066*** (0.0228)	0.1979*** (0.0216)
DailyRet_L1		-11.8576*** (1.3129)	-11.8539*** (1.3111)
DailyRet_L2		-7.1135*** (0.9567)	-7.0605*** (0.9480)
DailyRet_L3		-5.7773*** (0.8832)	-5.6490*** (0.8583)
DailyRet_L4		-3.1141*** (0.8703)	-2.9809*** (0.8680)
DailyRet_L5		-1.2782 (0.7764)	
R^2	0.6889	0.7053	0.7053
AIC	8917.11	8862.12	8862.51
BIC	8962.77	8927.36	8921.22

Next, normalized realized skewness (NRS), normalized realized kurtosis (NRK) and the Amihud illiquidity measure are incorporated into the HAR+Ret4 model. Note that NRS measures asymmetry, Amihud measures illiquidity, and NRK measures tail risk and jump intensity. The results in Table 3 show that the realized skewness is just marginally important with a p-value slightly higher than 0.05 and the Amihud illiquidity measure is significant at 5%. Worth noting is the realized kurtosis measure, which is very significant at 0.1% level, even after controlling for skewness and illiquidity. The data indicates that tail risk contains most of the incremental information for future volatility, while skewness and illiquidity contribute little once tail risk is included. This suggests NRK is the dominant realized moment. In terms of the model fit, among the realized-moment measures, normalized realized kurtosis (NRK) provides the largest improvement in model fit. Relative to the baseline HAR specification, the NRK produces the largest increase in adjusted R^2 and substantial reductions in both AIC and BIC. Although the specifications including NRS, NRK, and Amihud illiquidity achieve a marginally higher adjusted R^2 , the gain is economically negligible and comes at the cost of additional parameters. Moreover, the BIC favors the more parsimonious HAR-

NRK specification. Consequently, NRK is retained as the primary realized-moment variable in subsequent analysis.

Table 3: Augmented HAR model with higher moments and Amihud Illiquidity

This table reports the regression result of the augmented HAR-RV model with various higher realized moments and the illiquidity measure included. Coefficients are reported together with Newey-West standard error in parentheses. ***, **, and * denote significance at the 0.1%, 1% , 5% level, respectively. The parentheses by model indicate the equations specified in section 3.2.

Variable	Model (17) HAR+Ret4+NRS	Model (18) HAR+Ret4+NRK	Model (19) HAR+Ret4+Amihud	Model (20) HAR+Ret4+All
Intercept	-0.7145*** (0.0895)	-0.6768*** (0.0892)	-0.6664*** (0.0942)	-0.6150*** (0.0941)
HAR_D	0.2947*** (0.0229)	0.3230*** (0.0236)	0.2948*** (0.0224)	0.3172*** (0.0230)
HAR_W	0.4390*** (0.0303)	0.4105*** (0.0304)	0.4359*** (0.0301)	0.4172*** (0.0304)
HAR_M	0.1964*** (0.0217)	0.1928*** (0.0213)	0.2032*** (0.0214)	0.1970*** (0.0212)
DailyRet_L1	-11.1742*** (1.3888)	-11.1340*** (1.2473)	-11.1930*** (1.2969)	-10.6880*** (1.3163)
DailyRet_L2	-7.2005*** (0.9464)	-6.6918*** (0.9187)	-7.1280*** (0.9382)	-6.8795*** (0.9105)
DailyRet_L3	-5.6770*** (0.8589)	-5.3652*** (0.8738)	-5.7557*** (0.8570)	-5.4992*** (0.8693)
DailyRet_L4	-2.9858*** (0.8663)	-2.8425** (0.8732)	-3.0242*** (0.8662)	-2.8927*** (0.8693)
NRS_L1	-0.0225 (0.0121)			-0.0180 (0.0118)
NRK_L1		-0.0164*** (0.04034)		-0.0159*** (0.0034)
Amihud_scaled_L1			-1.9673* (0.8705)	-1.9100* (0.8785)
Adjusted R ²	0.7054	0.7069	0.7055	0.7073
AIC	8859.97	8834.70	8859.53	8831.04
BIC	8925.21	8899.94	8924.76	8909.31

Next, we consider augmenting the HAR model with implied variance derived from VIX and realized volatility of the VIX (VIXRV). Table 4 shows that the implied variance derived from VIX level is very significant. However, the realized volatility of the VIX index is not significant at all, indicating that it is only the level of VIX rather than the changes of VIX affecting the realized volatility. Thus, the VIX level captures almost all useful information, which is quite plausible economically. The VIX level already summarizes a huge amount of information about expected future volatility. The addition of VIXRV produces virtually no improvement in adjusted R^2 , while both AIC and BIC increase. In particular, the BIC favors the simpler HAR+Ret4+IV specification by more than seven points. Consequently, we retain the HAR+Ret4+IV model as the benchmark specification in subsequent analysis.

Table 4: Augmented HAR model with VIX components

This table reports the regression result of the augmented HAR model with the implied variance derived from VIX and realized volatility of VIX. Coefficients are reported together with Newey-West standard error in parentheses. ***, **, and * denote significance at the 0.1%, 1%, 5% level, respectively. The parentheses by model indicate the equations specified in section 3.2.

Variable	Model HAR+Ret4+IV (21a)	Model HAR+Ret4+IV+VIXRV (21b)
Intercept	-0.0563 (0.1040)	-0.1056 (0.1157)
HAR_D	0.2573 (0.0219)	0.2661*** (0.0233)
HAR_W	0.3225 (0.0283)	0.3228*** (0.0282)
HAR_M	0.0193 (0.0260)	0.01294 (0.0261)
DailyRet_L1	-8.8260 (1.0436)	-8.8642*** (1.0509)
DailyRet_L2	-5.1210 (0.8246)	-5.1303*** (0.8247)
DailyRet_L3	-4.1135 (0.8254)	-4.1283*** (0.8253)
DailyRet_L4	-2.2335** (0.8498)	-2.2454*** (0.8478)
IV_L1	0.4569*** (0.0353)	0.4554*** (0.0354)
VIXRV_L1		-0.0102 (0.0109)
Adjusted R ²	0.7200	0.7200
AIC	8605.72	8606.85
BIC	8670.96	8678.61

The next step is naturally to investigate if the effect of realized kurtosis remains even after controlling for the implied variance expected by the level of VIX index. As shown in Table 5, after controlling for the implied variance, we find that realized skewness is not significant at all for all the models considered, nor is the Amihud illiquidity measure. The small significance of liquidity effects we observed before largely disappears once option-implied information is included. However, realized kurtosis contains incremental predictive information about future realized volatility beyond what the level of VIX can provide. The specification including normalized realized kurtosis (NRK) provides the strongest performance. It achieves the highest adjusted R^2 among the parsimonious specifications and yields the lowest AIC and BIC. Adding realized skewness, Amihud illiquidity, or realized volatility of the VIX produces negligible improvements in model fit while increasing model complexity. Consequently, the HAR+Ret4+IV+NRK model is selected as the benchmark specification for subsequent analysis.

Table 5: augmented HAR-RV model with IV and other variables

This table reports the regression result of the augmented HAR-RV model with realized skewness, kurtosis, and Amihud illiquidity measure. Coefficients are reported together with Newey-West standard error in parentheses. ***, **, and * denote significance at the 0.1%, 1% , 5% level, respectively. The parentheses by model indicate the equations specified in section 3.2.

Variable	Model (22a) HAR+Ret4+ IV+NRS	Model (22b) HAR+Ret4+ IV+NRK	Model (22c) HAR+Ret4+ IV+Amihud	Model (22d) HAR+Ret4+ IV+All
Intercept	-0.0544 (0.1042)	-0.0249 (0.1052)	-0.0240 (0.1068)	0.0073 (0.1080)
HAR_D	0.2555*** (0.0218)	0.2803*** (0.0227)	0.2557*** (0.0214)	0.3081*** (0.0228)
HAR_W	0.3263*** (0.0283)	0.3032*** (0.0284)	0.3242*** (0.0281)	0.3081*** (0.0283)
HAR_M	0.0733 (0.0259)	0.0660 (0.0258)	0.0233 (0.0255)	0.0821 (0.0253)
DailyRet_L1	-8.3963*** (1.1281)	-8.2278*** (1.0222)	-8.8873*** (1.0402)	-7.9824*** (1.3163)
DailyRet_L2	-5.2200*** (0.8291)	-4.8194*** (0.8188)	-5.1708*** (0.8162)	-4.9460*** (0.8159)
DailyRet_L3	-4.1380*** (0.8260)	-3.8819*** (0.8212)	-4.1849*** (0.8174)	-3.9734*** (0.8123)
DailyRet_L4	-2.2400*** (0.8501)	-2.1208* (0.8553)	-2.2633** (0.8466)	-2.1563* (0.8523)
IV	0.4551*** (0.0353)	0.4513*** (0.0353)	0.4548*** (0.0355)	0.4479*** (0.0355)
NRS_L1	-0.0146 (0.0113)			-0.0180 (0.0111)
NRK_L1		-0.0146*** (0.0032)		-0.0142*** (0.0032)
Amihud_scaled_L1			-1.1818 (1.4199)	-1.1456 (1.4079)
Adjusted R ²	0.7200	0.7213	0.7200	0.7213
AIC	8605.71	8583.32	8605.84	8584.44
BIC	8677.47	8655.08	8677.60	8669.24

Next, we investigate if the tail-risk effect persists beyond one day by comparing three specifications that incorporate realized kurtosis: a one-lag model (NRK), a five-lag model (NRK5), and a weekly-average model (NRK_W). The one-lag model (NRK) is essentially the same model we present in Table 3, which is reported again for the convenience in Table 6. Table 6 shows that NRK5 model has both NRK_L1 and NRK_L3 significant at the 0.1% level. Therefore, tail-risk information affects future volatility over multiple days rather than only one day. We also observe a decaying pattern of the tail-risk effect on future volatilities, as the parameters become less negative from first lag to fifth lag. The joint Wald tests for the five-lag model were highly significant at the 0.1% significance level. It suggests that realized kurtosis has a persistent multi-day forecasting effect rather than a one-day anomaly. The five-lag specification delivers the highest adjusted R^2 and the lowest AIC among all models considered. Although the weekly-average specification yields a slightly lower BIC, the difference is modest. Taken together, the results suggest that tail-risk information persists over multiple days, motivating the use of the HAR+Ret4+IV+NRK5 model as the primary specification.

Table 6: Augmented HAR model with lagged Realized Kurtosis

This table reports the regression result of the augmented HAR model with the implied variance expected from VIX and multiple lagged realized kurtosis. Coefficients are reported together with Newey-West standard error in parentheses. ***, **, and * denote significance at the 0.1%, 1% , 5% level, respectively. The parentheses by the model indicate the equations specified in section 3.2.

Variable	Model (18) HAR+Ret4+IV+NRK	Model (23) HAR+Ret4+IV+NRK5	Model (24) HAR+Ret4+IV+NRK_W
Intercept	-0.0249 (0.1052)	-0.0915 (0. 0.10610)	-0.0013 (0.1062)
HAR_D	0.2803*** (0.0227)	0.2775 (0.0219)	*** 0.2603*** (0.0219)
HAR_W	0.3032*** (0.0284)	0.3082 (0.0285)	*** 0.3220*** (0.0282)
HAR_M	0.0660 (0.0258)	0.0125 (0.0255)	0.0114 (0.0256)
DailyRet_L1	-8.2278*** (1.0222)	-8.3332 (1.0157)	*** -8.7128*** (1.0229)
DailyRet_L2	-4.8194*** (0.8188)	-4.9918 (0.8167)	*** -4.9920*** (0.8225)
DailyRet_L3	-3.8819*** (0.8212)	-3.6543 (0.8253)	*** -3.9486*** (0.8266)
DailyRet_L4	-2.1208* (0.8553)	-2.0675* (0.8764)	-2.0406*** (0.8609)
IV_L1	0.4513*** (0.0353)	0.4485*** (0.0347)	0.4536*** (0.0347)
NRK_L1	-0.0146*** (0.0032)	-0.0140*** (0.0031)	
NRK_L2		0.0048 (0.0029)	
NRK_L3		-0.0143*** (0.0028)	
NRK_L4		-0.0032 (0.0028)	
NRK_L5		-0.0034 (0.0030)	
NRK_W			-0.0301*** (0.0067)
Adjusted R ²	0.7213	0.7228	0.7213
AIC	8583.32	8560.15	8583.09
BIC	8655.08	8658.00	8654.88

Table 7 reports the in-sample performance of the alternative HAR specifications. Augmenting the baseline HAR model with lagged returns substantially improves model fit. The adjusted R^2 increases from 0.6889 to 0.7053, while both AIC and BIC decline sharply. Similarly, augmenting the HAR+Ret4 model with implied variance also substantially improves model fit. The adjusted R^2 increases again from 0.7053 to 0.7200, with both AIC and BIC declining over 250 points. These results confirm that option-implied volatility contains valuable forward-looking information about future realized volatility. Among the realized moment measures considered, normalized realized kurtosis (NRK) provides the largest incremental improvement beyond the HAR+Ret4+IV benchmark. Adding NRK increases the adjusted R^2 to 0.7213 and further reduces both AIC and BIC. After incorporating multiple lagged NRK, the model fit improved again. We find that the HAR+Ret4+IV+NRK5 specification achieves the highest adjusted R^2 and the lowest AIC among all models considered. Although the NRK_W specification yields a slightly lower BIC (3 points lower) due to its more

parsimonious structure, the overall evidence still favors the NRK5 model. These results indicate that the predictive content of realized kurtosis extends over multiple trading days rather than being confined to the most recent observation.

Table 7: In-sample performance of the alternative HAR specifications

Models	Adjusted R ²	AIC	BIC
HAR+Ret4+IV+NRK5	0.7228	8560.15	8658.00
HAR+Ret4+IV+ NRK_W	0.7213	8583.09	8654.85
HAR+Ret4+IV+NRK	0.7213	8583.32	8655.08
HAR+Ret4+IV	0.7200	8605.71	8677.47
HAR+Ret4	0.7053	8862.51	8921.22
HAR	0.6889	9129.71	9162.34

Table 8 presents the out-of-sample forecasting results of the various HAR specifications. The HAR+Ret4+IV+NRK5 model delivers the lowest root mean square error (RMSE), mean absolute error (MAE), and mean squared error (MSE) among all competing specifications, indicating superior forecasting performance. Relative to the baseline HAR model, the inclusion of multiple lagged returns, implied variance and multiple lagged realized kurtosis produces meaningful improvements in forecast accuracy. Thus, the gains observed in-sample are largely preserved out-of-sample. This finding suggests that the predictive power of realized kurtosis is not a result of overfitting but instead reflects economically meaningful information about future volatility dynamics. The HAR+Ret4+IV+NRK^W specification also performs well and yields forecast accuracy close to that of the NRK5 model, providing additional evidence that persistent tail-risk information contributes to volatility forecasting. Consequently, the HAR+Ret4+IV+NRK5 specification emerges as the preferred forecasting model.

Table 8: Out-of-sample forecasting of the alternative HAR specifications

Models	RMSE	MAE	MSE
HAR+Ret4+IV+NRK5	0.6805	0.5138	0.4630
HAR+Ret4+IV+NRK_W	0.6814	0.5150	0.4643
HAR+Ret4+IV+NRK	0.6823	0.5154	0.4655
HAR+Ret4+IV	0.6836	0.5168	0.4673
HAR+Ret4	0.7016	0.5266	0.4923
HAR	0.7219	0.5412	0.5212

5. CONCLUSION

This paper shows that realized kurtosis contains persistent incremental information for forecasting realized volatility beyond the standard HAR framework with improvements demonstrated both in-sample and out-of-sample. The forecasting contribution of realized kurtosis remains economically and statistically significant even after controlling for volatility persistence and implied variance derived from option markets. Furthermore, the predictive effects of realized kurtosis persist across multiple trading days, indicating that tail-risk information is not fully incorporated into future volatility expectations immediately. In contrast, realized skewness and the Amihud illiquidity measure provide little additional forecasting information once realized kurtosis is included in the model. These findings suggest that tail-risk dynamics, rather than return asymmetry or temporary liquidity fluctuations, represent a key determinant of future volatility in the highly liquid ES futures market. Our findings have significant financial and economic implications. Economically, market shocks do not just disappear overnight, and consequently tail events affect volatility for multiple trading days. Thus, in the highly liquid ES futures market, tail risk is a more persistent and economically important predictor of future realized volatility than return asymmetry or temporary fluctuations in trading liquidity.

Our findings provide significant managerial implications for risk managers and option traders. Our results suggest that relying solely on historical volatility or option-implied volatility may underestimate the

persistence of market stress following extreme price movements. Risk managers may benefit from incorporating realized kurtosis into volatility forecasting models to improve stress testing and portfolio risk monitoring during turbulent market conditions. Likewise, option traders and volatility arbitrage strategies may benefit from incorporating realized kurtosis when forecasting future realized volatility.

Our research also has several limitations. First, the empirical analysis focuses exclusively on the S&P 500 E-mini futures market. Although the ES contract is among the world's most liquid financial instruments, the results may not generalize to less liquid assets. Second, this study adopts the HAR framework because of its simplicity and strong forecasting performance. More flexible nonlinear or machine learning models may capture additional dynamic relationships that are beyond the scope of this study. Finally, changes in the global economic and geopolitical environment—including evolving trade, tariff, and tax policies, as well as international conflicts—may alter market volatility, liquidity conditions, and investor behavior in ways that differ from the historical sample examined in this study.

Several directions for future research remain. First, it would be useful to examine whether the forecasting power of realized kurtosis extends to other asset classes, including individual equities, fixed-income securities, commodities, and cryptocurrencies. Second, future studies could investigate whether the predictive role of realized kurtosis varies across different market regimes, particularly during periods of financial stress. Finally, further research may explore whether liquidity measures play a more important role in forecasting volatility for less liquid assets and markets.

REFERENCES:

- Amaya, D., Christoffersen, P., Jacobs, K. and Vasquez, A., "Does realized skewness and kurtosis predict the cross-section of equity returns?" *Journal of Financial Economics*, Volume 118, Issue 1, Pages 135–167, 2015.
- Amihud Y., "Illiquidity and stock returns: cross-section and time-series effects", *Journal of Financial Markets*, Volume 5, Number 1, Pages 31-56, 2002.
- Andersen, T. G., and Bollerslev, T., "Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependences", *Journal of Finance*, Volume 53, Number 1, Pages 219-265, 1998.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P., "Modeling and forecasting realized volatility", *Econometrica*, Volume 71, Number 2, Pages 579-625, 2003.
- Bae, K. and Lee, S., "Realized higher-order comoments", *Quantitative Finance*, Volume 21, Number 3, Pages 421-429, 2021.
- Barndorff-Nielsen, O.E. and Shephard, N., "Power and Bipower Variation with Stochastic Volatility and Jumps", *Journal of Financial Economics*, Volume 2, Number 1, Pages 1-37, 2004.
- Bonato, M., Cepni, O., Gupta, R. and Pierdeioch, C., "Forecasting Realized Volatility of International REITs: the Role of Realized Skewness and Realized Kurtosis", *Journal of Forecasting*, Volume 41, Pages 303-315, 2021.
- Corsi, F., "A Simple Approximate Long-memory Model of Realized Volatility", *Journal of Financial Econometrics*, Volume 7, Number 2, Pages 174-196, 2009.
- Huang, X., and Tauchen G., "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, Volume 3, Number 4, Pages 456-499, 2005.
- Kozhan, R., Neuberger, A. & Schneider, P., "The Skew Risk Premium in the Equity Index Market", *The Review of Financial Studies*, Volume 26, Number 9, Pages 2174-2203, 2013.

Liang, C., Wei, Y. and Zhang, Y., "Is Implied Volatility More Informative for Forecasting Realized Volatility: An International Perspective", *Journal of Forecasting*, Volume 39, Pages 1253-1276, 2020.

Mei, D., Liu, J., Ma, F., and Chen, W., "Forecasting stock market volatility: Do realized skewness and kurtosis help?" *Physica A: Statistical Mechanics and its Applications*, Volume 481, September, Pages 153-159, 2017.

Neuberger, A., "Realized Skewness", *The Review of Financial Studies*, Volume 25, Number 11, Pages 3423-3455, 2012.

Neuberger, A. & Payne, R. G. "The Skewness of the Stock Market at Long Horizons", *The Review of Financial Studies*, Volume 34, Number 3, Pages 1572-1616, 2021.

AUTHOR PROFILE:

Dr. Megan Y. Sun earned her Ph.D. from Kent State University in 2003. Currently she is a professor of Finance at University of Wisconsin – River Falls.

Dr. Dawn Hukai earned her Ph.D. from University of Minnesota in 1998. Currently she is a professor of Accounting at University of Wisconsin – River Falls.

DYNAMIC CAPABILITIES AND RESOURCE RECONFIGURATION IN THE JAPANESE EV INDUSTRY: THE CASE OF TOYOTA AND HONDA

Woonghee Lee, Hanyang University, Seoul, South Korea

dx.doi.org/10.18374/JABE-26-3.7

ABSTRACT

This study examines how established firms reconfigure resources and capabilities under technological uncertainty by comparing Toyota and Honda's electric vehicle (EV) strategies. Drawing on Teece's (2007) dynamic capabilities framework of sensing, seizing, and transforming, the study adopts a comparative longitudinal case approach covering 2019–2026. Despite facing broadly similar environmental conditions, the two firms pursued contrasting strategies. Toyota followed a multi-pathway approach characterized by gradual capability accumulation, resource recombination, and preservation of technological options. Honda initially pursued a more concentrated EV strategy involving substantial EV-specific investments and external partnerships, but subsequently reassessed and redeployed resources as market conditions changed. The findings suggest that existing resource configurations influence how firms sense technological change and make resource commitments, while these commitments subsequently shape their capacity for further reconfiguration. The study highlights the recursive relationship between dynamic capabilities and resource configuration under technological uncertainty.

Keywords: Dynamic Capability, EV Industry, Toyota, Honda

1. INTRODUCTION

The electric vehicle (EV) industry is characterized by a high degree of technological uncertainty and frequent changes in market and regulatory conditions. This paper examines how two major Japanese automobile manufacturers, Toyota and Honda, have responded differently to the transition toward electric vehicles despite facing broadly similar environmental conditions. The two firms provide a particularly suitable basis for comparative analysis. Both are long-established Japanese automobile manufacturers with extensive international operations, strong engineering and manufacturing capabilities, globally recognized brands, and substantial experience in vehicle electrification. Despite these similarities, however, Toyota and Honda have differed in the timing, direction, and degree of their commitments to EVs and in the ways in which they have developed, acquired, and deployed resources in response to the transition.

The comparison of Toyota and Honda therefore provides an opportunity to address a broader strategic question: why and how do established firms facing similar environmental changes differ in the ways they reconfigure their resources and capabilities in response to technological uncertainty? To address this question, this paper adopts the dynamic capabilities perspective as its primary theoretical framework (Teece, Pisano, and Shuen, 1997). In particular, the analysis draws on Teece's (2007) framework of sensing, seizing, and transforming to examine how Toyota and Honda have interpreted changes in the EV environment, made strategic commitments in response to perceived opportunities and threats, and subsequently reconfigured their existing and newly acquired resources and capabilities.

The study adopts a comparative longitudinal case approach covering the period from 2019 to 2026 and is based primarily on archival data. Primary archival sources include Toyota's and Honda's annual and integrated reports, strategy briefings, press releases, investor presentations, and other corporate disclosures. These materials provide information on the firms' strategic objectives, technological priorities, investment decisions, partnerships, and organizational changes. To complement and triangulate company-provided information, secondary data were collected from major international and Japanese business publications, including Nikkei Asia, The Wall Street Journal, Bloomberg, and the Financial Times. Government and industry sources, including publications and statistics from Japan's Ministry of Economy,

Trade and Industry (METI), Ministry of Land, Infrastructure, Transport and Tourism (MLIT), and Japan Automobile Manufacturers Association (JAMA), were also consulted to establish the broader institutional and market context. The use of multiple independent sources allows major strategic decisions and changes in corporate direction to be cross-checked and traced over time, thereby providing a longitudinal basis for comparing the two firms' processes of resource reconfiguration.

Through this comparative analysis, the paper seeks to contribute to the dynamic capabilities literature by examining how differences in firms' existing resource configurations and strategic responses may generate distinct pathways of resource reconfiguration even when firms confront broadly similar environmental conditions (Eisenhardt and Martin, 2000). More specifically, the Toyota–Honda comparison provides an opportunity to examine how inherited resources may influence the ways in which firms sense technological change and seize emerging opportunities, while the resulting strategic commitments subsequently shape their ability to transform and reconfigure their resource bases.

2. PRIOR RESEARCH ON DYNAMIC CAPABILITY AND AUTOMOBILE INDUSTRY

Building on, but extending, the resource-based view of the firm, Teece, Pisano, and Shuen (1997) define dynamic capabilities as a firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments. Subsequent research has further elaborated the processes through which such adaptation occurs. Eisenhardt and Martin (2000) conceptualize dynamic capabilities as identifiable organizational and strategic processes, such as product development, strategic decision making, alliance formation, and resource allocation, through which firms integrate, acquire, and reconfigure resources. Helfat and Peteraf (2003) emphasize that organizational capabilities evolve through different stages of development, maturity, transformation, and decline.

Teece (2007) further develops the dynamic capabilities perspective by disaggregating dynamic capabilities into three fundamental clusters: sensing, seizing, and transforming. In rapidly changing environments, possessing valuable resources alone is insufficient; firms must continually create, extend, upgrade, protect, and reconfigure their asset bases. Sensing refers to identifying and interpreting technological developments, changing customer needs, competitive threats, and emerging opportunities. It involves processes of scanning, searching, learning, and interpreting developments both within and outside the firm's existing technological and market domains. Firms exposed to similar environmental signals may nevertheless interpret their significance differently because managerial cognition, prior experience, and existing technological positions affect how opportunities and threats are perceived.

Once an opportunity or threat has been sensed, the firm must seize it through strategic choices and resource commitments. Seizing involves decisions concerning investments in technologies, products, business models, organizational arrangements, and complementary assets. Such decisions become particularly difficult under technological uncertainty because several competing technological trajectories may coexist and their future commercial viability may remain unclear. Firms must therefore decide not only which opportunities to pursue, but also when and how extensively to commit their resources. Teece (2007) notes that firms with strong positions in complementary assets may sometimes possess greater ability to delay commitment, while firms in less favorable positions may have stronger incentives to act rapidly. Consequently, firms that sense essentially the same technological transition may make substantially different strategic commitments.

Finally, transforming involves the continuous renewal and reconfiguration of the firm's resource and capability base as technologies and markets evolve. Existing resources can both facilitate and constrain these processes. Technological knowledge, manufacturing capabilities, complementary assets, and organizational experience provide foundations for adaptation, while path-dependent routines and commitments to established technologies may inhibit radical transformation. Teece (2007) observes that established capabilities, complementary assets, and administrative routines can bias incumbent organizations against radical innovation even after emerging opportunities have been recognized.

The dynamic capabilities perspective has increasingly been applied to the transformation of the automobile industry. Teece (2019), for example, analyzes the emergence of electric, connected, shared, and autonomous mobility as simultaneous paradigm shifts and argues that sensing and sensemaking capabilities are particularly important under the resulting deep uncertainty. His analysis also emphasizes the importance of complementary assets and the capacity of both incumbent and emerging Chinese firms to develop new capabilities. Research on incumbent automakers similarly finds that the transformation requires multiple dynamic capabilities for determining and executing new strategies rather than merely acquiring individual new technologies (Hoeft, 2020). Knobbe and Proff (2020), examining automotive suppliers under digitalization, identify different patterns through which firms activate sensing, seizing, and reconfiguration capabilities, suggesting substantial heterogeneity in adaptation even within the same industry.

Research specifically addressing electrification provides further support for this view. Loder, Rinscheid, and Wüstenhagen's (2024) comparative study of German automobile manufacturers finds substantial differences in how incumbent firms perceive the opportunities and risks associated with electric mobility. Their analysis connects these differences in sensing capabilities to managerial cognitive frames, prior technological experience, network embeddedness, and leadership. Collectively, these studies demonstrate the usefulness of dynamic capabilities for explaining automobile-industry transformation, while also indicating that incumbent firms exposed to the same technological disruption can differ considerably in how they interpret change and mobilize resources in response. However, dynamic capability view, which was applied in German and Chinese auto industry, has not been adopted in Japanese automobile industry yet.

The sensing–seizing–transforming framework may provide an appropriate theoretical lens for comparing Toyota and Honda's responses to the EV transition as it was applied in other national settings such as Germany and China (Loder, Rinscheid & Wüstenhagen, 2024; Teece, 2007). The analysis that follows examines how the two firms have sensed changes associated with electrification, seized emerging opportunities through different technological choices, investments, and partnerships, and transformed their resource bases through the development, acquisition, combination, and redeployment of resources and capabilities. The comparison thereby seeks to explain why two incumbent automobile manufacturers operating under broadly similar environmental conditions may develop different pathways of resource reconfiguration.

3. THE EV TRANSITION IN TOYOTA AND HONDA

Toyota and Honda entered the recent transition toward electric vehicles from broadly similar positions as established Japanese automobile manufacturers with strong technological, manufacturing, and brand resources. Both firms had accumulated substantial experience in vehicle electrification, particularly through hybrid electric vehicles (HEVs), before the rapid expansion of the global battery electric vehicle (BEV) market. They also faced a common Japanese policy environment that promoted electrification while allowing multiple technological pathways: the Japanese government's objective of achieving 100 percent electrified new passenger-car sales by 2035 includes HEVs, PHEVs, BEVs, and fuel-cell vehicles rather than requiring an exclusively BEV transition (Ministry of Economy, Trade and Industry [METI], 2023). Despite these common conditions, Toyota and Honda differed considerably in how they interpreted the pace and direction of the EV transition and, consequently, in the timing and nature of their resource commitments.

3.1 Toyota: A Multi-Pathway Approach to Electrification

Toyota entered the EV transition with an unusually strong legacy position in hybrid technology, developed through more than two decades of experience following the introduction of the Prius in 1997. Rather than treating BEVs as the only viable route toward carbon neutrality, Toyota maintained what it termed a “multi-pathway” strategy encompassing HEVs, PHEVs, BEVs, fuel-cell vehicles, and other powertrain technologies. Independent reporting consistently characterized Toyota's approach as more cautious toward an EV-only transition than that of many global competitors (Reuters, 2024a, 2024b).

This position did not, however, imply an absence of substantial investment in BEVs. Toyota sharply increased its stated BEV ambitions in December 2021, announcing a target of 30 BEV models and 3.5 million annual BEV sales globally by 2030, accompanied by substantial planned investment in electrification and batteries (Toyota Motor Corporation, 2021). The announcement represented a significant increase from Toyota's earlier plans, while retaining HEVs, PHEVs, and fuel-cell vehicles within the broader electrification portfolio. Thus, Toyota's increasing commitment to BEVs was embedded within, rather than substituted for, its existing multi-technology strategy.

A further acceleration occurred after Koji Sato became president in 2023. Toyota created a dedicated BEV Factory organization and announced next-generation BEVs to be introduced from 2026. According to Reuters, the new unit was expected to account for approximately 1.7 million vehicles by 2030, or roughly half of Toyota's previously announced 3.5-million-unit BEV target (Reuters, 2023a). The company simultaneously announced technological changes in vehicle design, manufacturing processes, and batteries intended to improve the competitiveness of its next-generation BEVs. These developments suggest that Toyota was progressively establishing organizational and technological resources specifically oriented toward BEVs even while preserving its broader powertrain portfolio.

Battery technology became a particularly important area of resource development. Toyota pursued multiple battery technologies intended for different vehicle segments and price-performance requirements, including lithium-ion, lower-cost bipolar lithium iron phosphate, and all-solid-state batteries. Of particular relevance from a resource-reconfiguration perspective, Toyota sought to transfer bipolar battery technology already developed for hybrid vehicles into its future BEVs (Toyota Motor Corporation, 2023a). The company also partnered with Idemitsu Kosan to develop mass-production technology and supply capabilities for solid electrolytes used in all-solid-state batteries (Toyota Motor Corporation & Idemitsu Kosan, 2023). These initiatives illustrate Toyota's combination of internal capability development, recombination of existing electrification knowledge, and selective acquisition of complementary external capabilities.

Toyota also progressively expanded its battery manufacturing resources in North America. Its North Carolina battery facility, initially announced in 2021, received successive increases in investment as Toyota expanded its electrification plans (Reuters, 2021; Reuters, 2023b). By October 2023, Toyota had announced an additional \$8 billion investment, bringing planned investment in the facility to approximately \$13.9 billion. Importantly, four of the plant's six planned production lines were intended to support hybrid vehicles and two to support BEVs. This configuration is significant because the investment strengthened Toyota's battery capabilities without making the entire facility dependent on a rapid transition to BEVs. The plant began production in 2025 as Toyota simultaneously expanded hybrid production in North America (Reuters, 2025).

Developments in the global EV market after 2023 increased the strategic importance of this flexibility. As the pace of BEV adoption moderated in some major markets, Toyota benefited from strong demand for hybrids. Reuters reported in early 2024 that Toyota's multi-pathway strategy, previously criticized for being too cautious toward BEVs, was benefiting from consumer concerns about EV prices, driving range, and charging infrastructure (Reuters, 2024a). By mid-2024, hybrids accounted for 37 percent of Toyota's sales, compared with only 9 percent in 2018, and Toyota was increasingly moving major models toward hybrid-only offerings in the U.S. market (Reuters, 2024c). The Financial Times similarly characterized Toyota's continued investment in efficient combustion engines for hybrid applications as part of a broader diversified approach at a time when EV demand was developing less rapidly than some manufacturers had anticipated (Financial Times, 2024).

Toyota's EV transition can therefore be characterized as progressive resource accumulation combined with technological portfolio preservation. Rather than replacing its existing hybrid and manufacturing capabilities with an exclusively BEV-oriented resource base, Toyota continued exploiting these established resources while developing additional capabilities in BEVs, batteries, software, and new manufacturing technologies. Its approach consequently left a relatively large proportion of its resource base applicable across alternative powertrain trajectories.

3.2 Honda: Accelerated Commitment and Subsequent Reassessment

Honda entered the same transition with substantial engineering and hybrid capabilities but initially articulated a more concentrated long-term commitment to BEVs. In 2022, Honda announced plans to launch 30 EV models globally and achieve annual EV production of more than two million units by 2030, while allocating approximately ¥5 trillion over ten years to electrification and software technologies (Honda Motor Co., 2022a). This commitment formed part of Honda's longer-term objective of making BEVs and fuel-cell electric vehicles account for 100 percent of global automobile sales by 2040.

A notable feature of Honda's initial approach was the extensive use of external partnerships to accelerate access to capabilities it did not intend to develop entirely in-house. Honda expanded its relationship with General Motors in 2022 to co-develop affordable EVs based on a new architecture using GM's Ultium battery technology, with mass-market vehicles initially expected from 2027 (Reuters, 2022a). The collaboration complemented Honda's earlier arrangement under which GM would produce the Honda Prologue and an Acura EV using its EV technology. The proposed mass-market collaboration was subsequently abandoned in 2023 as the partners concluded that the business case had become difficult, illustrating the uncertainty surrounding the economics of affordable EV development (Reuters, 2025a).

Honda simultaneously invested heavily in battery capabilities and supply relationships. Its North American strategy included a battery joint venture with LG Energy Solution and the development of an EV manufacturing hub in Ohio. Rather than relying exclusively on external battery procurement, Honda sought to establish dedicated regional production capacity closely integrated with its future EV manufacturing system. These commitments were part of a broader attempt to construct an EV-specific resource base rather than merely add individual electric models to its existing product portfolio.

Honda also looked beyond the traditional automobile industry for capabilities associated with increasingly software-defined vehicles. In 2022, Honda and Sony established Sony Honda Mobility, combining Honda's vehicle engineering and manufacturing capabilities with Sony's strengths in software, entertainment, sensing, and digital technologies. Independent reporting described the venture as an effort to combine Honda's automotive expertise with Sony's software and gaming capabilities to compete with emerging EV rivals. The resulting AFEELA project therefore represented not simply a new EV model but an attempt to construct a new configuration of complementary automotive and digital resources.

Honda's EV commitment intensified considerably in 2024. The company announced that it would approximately double planned investment in electrification and software to ¥10 trillion through fiscal 2031. The strategy included a new generation of Honda 0 Series EVs, greater software-development capability, lower battery procurement costs, and more integrated EV production processes. Honda planned seven models in the 0 Series by 2030 and sought to reduce North American battery procurement costs by more than 20 percent and production costs by approximately 35 percent. These commitments indicate that Honda's strategy had moved beyond experimental EV development toward a substantial reconfiguration of its technological and manufacturing resource base.

The most visible manifestation of this commitment was Honda's proposed comprehensive EV value chain in Ontario, Canada. In April 2024, Honda announced plans for approximately C\$15 billion in investment, including investments by joint-venture partners, covering EV assembly, battery production, cathode materials, and separators. Reuters characterized the project as Honda's largest-ever investment in Canada and reported planned annual capacity of approximately 240,000 EVs and 36 GWh of batteries. The project represented an effort to develop an integrated regional EV resource system rather than depend extensively on imported batteries and components. In contrast to Toyota's North Carolina battery investment, much of which supported multiple electrified powertrains, Honda's Canadian initiative was explicitly structured around an anticipated large-scale expansion of BEVs.

By 2025, however, Honda's assessment of the market had changed substantially. Slower-than-anticipated EV demand, changes in government policies, and continued uncertainty over the speed of electrification led the company to reduce planned electrification and software investment by approximately 30 percent,

from ¥10 trillion to ¥7 trillion. Honda also abandoned its expectation that EVs would represent 30 percent of automobile sales in 2030 and increased its emphasis on hybrids, planning 13 new hybrid models and targeting substantially higher hybrid sales by the end of the decade (Reuters, 2025b). The C\$15 billion Canadian EV project was postponed for approximately two years, explicitly reflecting the slowdown in the EV market (Associated Press, 2025).

Honda's reassessment became still more pronounced in 2026. A broader reversal of EV commitments among global incumbent manufacturers produced large industry-wide write-downs as automakers responded to weaker-than-expected EV demand in the United States and Europe, changing regulatory support, and intense competition from Chinese manufacturers. Honda announced restructuring charges that could reach approximately ¥2.5 trillion as it scaled back its EV strategy and shifted greater attention toward hybrids (Reuters, 2026a). The adjustment was therefore not an isolated product decision but represented a significant reconsideration of resources previously committed to an anticipated EV-centered transition.

The strategic reversal also affected resources organized outside Honda's formal boundaries. In March 2026, Sony Honda Mobility discontinued development and planned launch of the AFEELA 1 and its second model after Honda's EV-strategy reassessment meant that technologies and assets originally expected from Honda would no longer be available as planned (Sony Group Corporation, Honda Motor Co., & Sony Honda Mobility, 2026). Reuters subsequently reported that the Sony-Honda venture would scale down its operations (Reuters, 2026b). The episode illustrates how strategic reconfiguration can become more complicated when resources and capabilities are distributed across organizational boundaries and jointly committed to a particular technological trajectory.

Honda's trajectory can therefore be characterized as accelerated resource commitment followed by substantial reassessment and redeployment. The firm initially sought to construct a relatively EV-specific resource configuration through internal development, major manufacturing investments, and extensive external partnerships. When its expectations concerning the speed and economics of the EV transition changed, Honda demonstrated a willingness to reverse some previous commitments, absorb substantial adjustment costs, and redirect resources toward hybrids and other technologies. The contrast with Toyota is thus not simply one of greater versus lesser commitment to EVs, but of different patterns in the timing, specificity, and reversibility of resource commitments under technological uncertainty.

4. COMPARATIVE ANALYSIS OF DYNAMIC CAPABILITIES AND RESOURCE RECONFIGURATION

The different trajectories of Toyota and Honda can be interpreted through Teece's (2007) distinction among sensing, seizing, and transforming capabilities. Both firms recognized electrification as a major technological transition and invested substantially in new technologies. Their differences lie less in whether they recognized the EV opportunity than in how they interpreted uncertainty, how strongly and when they committed resources, and how they structured the transformation of their existing resource bases.

4.1 Sensing: Different Interpretations of the Same Technological Transition

Sensing involves more than detecting the existence of a technological change. It requires managers to interpret its direction, timing, market potential, and relationship to other technological trajectories. Teece (2007) emphasizes that firms facing the same information can interpret opportunities differently because sensing involves managerial judgment under uncertainty.

Both Toyota and Honda clearly sensed electrification as an important long-term development. The principal difference concerned their assessment of the speed and technological concentration of the transition. Toyota consistently interpreted electrification as a technologically plural process. Its strategy treated BEVs as increasingly important but retained HEVs, PHEVs, and fuel cells as alternative means of achieving carbon reduction under heterogeneous infrastructure and market conditions. Even while announcing ambitious BEV investments, Toyota emphasized consumer choice and uncertainty over which technologies would dominate in different markets.

Honda's sensing process led to a stronger initial conclusion about the future centrality of BEVs. Its commitment to 100 percent BEV/FCEV automobile sales by 2040, plans for more than two million EVs annually by 2030, and large-scale EV-specific manufacturing and battery investments indicate that Honda interpreted the transition as requiring a relatively rapid strategic repositioning. The subsequent EV slowdown, regulatory changes, and intensified Chinese competition forced Honda to revise this interpretation. By 2025 and particularly 2026, Honda explicitly acknowledged that assumptions underlying its earlier EV strategy had changed. Its decision to reduce EV investment, expand hybrids, and cancel three major North American EV programs can therefore be viewed as an act of re-sensing rather than simply strategic retreat.

The comparison suggests that sensing capabilities should not be evaluated solely according to how rapidly firms recognize a technological opportunity. Under high uncertainty, sensing also involves judging the pace of transition and preserving the ability to update initial interpretations as new information emerges.

Proposition 1. Under conditions of high technological uncertainty, firms' existing resource configurations influence not only which technological opportunities they sense, but also their assessment of the timing and pace with which those opportunities should be pursued.

This proposition extends the role of existing resources beyond providing inputs for strategic action. Toyota's large hybrid resource base made continued exploitation of multiple powertrain technologies economically and technologically viable and may therefore have reinforced a more gradual interpretation of the BEV transition. Honda's different resource and strategic position contributed to a stronger commitment to constructing a new EV-centered configuration.

4.2 Seizing: Portfolio Preservation versus Concentrated Commitment

The largest difference between Toyota and Honda becomes visible in seizing. Once opportunities are recognized, seizing concerns decisions about technologies, investment timing, product architectures, alliances, and complementary assets.

Toyota's resource commitments were substantial but comparatively diversified. The firm expanded BEV models, created a dedicated BEV organization, developed next-generation software and batteries, and invested heavily in battery production. At the same time, it continued expanding hybrid systems and developing fuel-cell technologies. Many investments, particularly in battery manufacturing, could support multiple electrified powertrains. Toyota's North Carolina battery plant, for example, was designed to supply HEVs, PHEVs, and BEVs rather than being dependent exclusively on rapid BEV growth.

This strategy can be interpreted as option-preserving resource commitment. Toyota committed enough resources to develop significant BEV capabilities but retained resources that would remain economically useful if the transition proceeded more slowly or differently than expected. Its large hybrid business also generated continuing returns while BEV technology and demand evolved.

Honda's seizing process initially involved relatively more EV-specific commitments. Honda established dedicated EV platforms, the Honda 0 Series, an EV Hub, battery joint ventures, plans for a vertically integrated Canadian EV value chain, and multiple partnerships designed specifically to accelerate EV and software capability acquisition. This strategy potentially allowed Honda to accelerate the closing of capability gaps. Partnerships with GM, LG Energy Solution, Sony, CATL, GS Yuasa, and other firms gave Honda access to resources that would have been costly and time-consuming to develop independently. In dynamic-capability terms, Honda actively orchestrated external resources rather than limiting transformation to internal capability development.

However, concentrated commitment also increased exposure to changes in the expected technological trajectory. When EV growth slowed, some dedicated resources became difficult to redeploy without substantial costs. Honda's 2026 cancellation of three North American EV models and subsequent

restructuring illustrates the potential downside of early commitment under uncertainty. This leads to a second proposition:

Proposition 2. Under high technological and market uncertainty, resource commitments that preserve applicability across multiple technological trajectories provide greater strategic adaptability than highly technology-specific commitments, although potentially at the cost of slower initial specialization.

This proposition does not imply that diversified commitment is universally superior. If BEV adoption had accelerated more rapidly, Honda's earlier concentrated investments might have generated first-mover and learning advantages, while Toyota's cautious portfolio approach could have produced competitive delay. The relative value of the two approaches therefore depends partly on how uncertainty is eventually resolved.

4.3 Transforming: Recombination versus Substitution and Reversal

The cases also reveal different patterns of transforming. Both companies developed new resources, but they differed in the relationship between new capabilities and their legacy resource bases. Toyota's transformation largely involved resource recombination. Rather than replacing its established electrification, manufacturing, and supply-chain capabilities, Toyota attempted to combine them with new BEV, battery, and software capabilities. Knowledge developed through hybrids—battery durability, electric motors, power electronics, energy management, and manufacturing—could partly be redeployed toward BEVs. Similarly, Toyota adapted bipolar battery technologies previously used in hybrids for lower-cost BEV batteries.

The result was an evolutionary form of transformation in which established resources continued to generate value while new resources accumulated around them. Toyota's organization did undergo structural change, including creation of BEV Factory and stronger integration with Woven by Toyota, but the new organization remained embedded within a broader multi-powertrain strategy.

Honda's transformation involved a larger element of resource augmentation and substitution. Because it sought to move relatively rapidly toward a BEV-centered future, Honda obtained capabilities externally where its existing resource configuration appeared insufficient. The GM relationship supplied an initial EV platform; LG Energy Solution supplied battery production capabilities; Sony contributed digital and sensing technologies; and other partnerships supplemented battery materials and software capabilities. External resource orchestration can be an effective dynamic capability. Teece (2007) explicitly recognizes that firms operating in rapidly changing technological environments must integrate internally and externally generated innovations. However, extensive reliance on cospecialized external relationships may also increase reconfiguration complexity when strategy changes. The adjustment of Sony Honda Mobility following Honda's 2026 reassessment demonstrates that changes to one firm's resource commitments can alter the value and viability of resources organized across firm boundaries.

Honda's subsequent shift back toward HEVs demonstrates another important feature of transforming capabilities: transformation is not necessarily unidirectional. Dynamic capability may require reversing previous resource commitments, accepting sunk costs, and redeploying assets when new information changes the expected value of alternative strategies. Honda's decision to convert part of its joint-venture battery capacity toward hybrid batteries is a concrete example of physical asset redeployment following strategic re-sensing. This suggests:

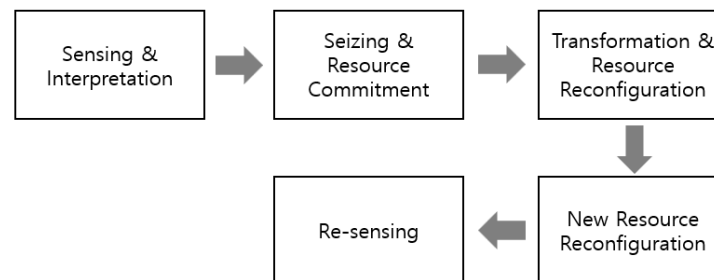
Proposition 3. The greater the specificity and interdependence of resources created for an emerging technological trajectory, the greater the cost and organizational complexity of subsequent resource reconfiguration when that trajectory changes.

This proposition connects dynamic capabilities with the characteristics of the resources being transformed. The ability to transform is not independent of prior seizing decisions: strategic commitments made during one period shape both the opportunities and constraints facing future transformation.

4.4 Dynamic Capabilities as a Recursive Process

Taken together, the Toyota and Honda cases suggest that sensing, seizing, and transforming should not necessarily be understood as a simple linear sequence. Instead, they form a recursive process. The process can be represented in <Figure 1> below

<Figure 1> Existing Resource Configuration



Toyota's experience illustrates relatively continuous iteration within this cycle. Its resource configuration supported several electrification trajectories; changing market information altered the allocation of resources among them without requiring a fundamental reversal of the overall strategic architecture.

Honda's experience exhibits a more pronounced cycle. Initial sensing generated a strong EV commitment; seizing produced substantial EV-specific resources and partnerships; subsequent changes in demand, regulation, and competition generated re-sensing; and the firm then began transforming the previously constructed resource configuration through cancellations, redeployment, and renewed investment in hybrids.

The comparison therefore highlights an important distinction between possessing resources that enable adaptation and configuring resources in ways that preserve future adaptability. Dynamic capabilities concern not merely the speed with which firms change, but also the design of resource commitments that allow further change when environmental uncertainty remains unresolved.

5. DISCUSSION AND CONCLUSION

This study has examined how Toyota and Honda, two incumbent Japanese automobile manufacturers exposed to broadly similar technological and institutional conditions, adopted different approaches to resource reconfiguration during the transition toward electric vehicles. Using Teece's (2007) sensing–seizing–transforming framework, the analysis suggests that differences between the two firms emerged not because one recognized electrification while the other failed to do so, but because they interpreted the timing, pace, and technological concentration of the transition differently and consequently made different resource commitments.

Toyota pursued a multi-pathway strategy in which existing hybrid, manufacturing, and supply-chain capabilities remained central while new BEV, battery, and software capabilities were progressively added. Honda adopted a relatively more concentrated approach to BEV transformation, investing heavily in EV-specific products, manufacturing systems, battery value chains, software, and external partnerships. When the expected pace of EV adoption changed, Honda subsequently undertook a significant reassessment and reallocation of resources toward hybrids while reducing several earlier EV commitments.

The comparison provides several implications for the dynamic capabilities literature. First, existing resource configurations may influence the exercise of dynamic capabilities at an earlier stage than is sometimes assumed. Resources are commonly viewed as the objects that dynamic capabilities integrate or

reconfigure. The Toyota–Honda comparison suggests that existing resources may also shape managerial sensing itself. Firms possessing different technological knowledge, complementary assets, market positions, and organizational histories may interpret the same external technological signals differently. Resource configuration is therefore both an antecedent and an outcome of dynamic capabilities.

Second, the cases draw attention to strategic flexibility in the seizing process. Seizing is commonly associated with commitment to promising opportunities. Under deep technological uncertainty, however, an equally important managerial capability may be determining the appropriate degree of commitment. Toyota's multi-pathway approach can be interpreted as maintaining technological options while accumulating BEV capabilities. Honda's experience demonstrates the potential advantages and costs of stronger commitment: early specialization can accelerate learning and capability building, but highly specific resource investments become more costly to reverse when technological or regulatory expectations change.

Third, the comparison suggests that transforming includes both forward capability creation and the reversal or redeployment of previous commitments. Honda's reduction of EV investments, cancellation of models, restructuring of partnerships, and conversion of part of planned EV battery resources toward hybrids are not outside the dynamic-capabilities framework. Rather, they illustrate the difficult process of reconfiguration after earlier strategic assumptions change. Strong dynamic capabilities should therefore not be equated with unwavering commitment to an initial technological strategy.

These observations lead to a broader implication. In environments characterized by high technological uncertainty, the effectiveness of dynamic capabilities may depend partly on whether firms create resource configurations that maintain what may be called reconfiguration flexibility - the capacity to alter the direction, timing, and intensity of strategic commitments without destroying excessive amounts of previously created value. Resources that are useful across several technological trajectories may provide firms with greater adaptability, whereas highly cospecialized resources can generate greater performance if the anticipated trajectory materializes but larger adjustment costs if it does not.

The Toyota and Honda cases also have managerial implications for other incumbent manufacturers confronting technological disruption. Managers need not choose simply between defending legacy technologies and abandoning them in favor of emerging technologies. Existing capabilities may serve as complementary resources for developing new businesses and may provide financing, technological knowledge, manufacturing expertise, or strategic flexibility during periods of transition. At the same time, excessive reliance on successful legacy resources can create inertia, as Teece (2007) warns. The managerial challenge is therefore to distinguish resources that constrain adaptation from resources that can be recombined to support it.

This study is subject to several limitations. The analysis is based on two firms from a single national setting and therefore cannot establish whether the observed patterns apply broadly across industries or countries. Moreover, the EV transition remains ongoing. Toyota's multi-pathway strategy may appear well suited to the current period of uneven EV adoption, but a rapid future acceleration of BEVs could alter the relative advantages of the two strategies. Similarly, Honda's recent reassessment is too recent to evaluate its long-term consequences. The propositions developed here should therefore be regarded as theory-building arguments that require further examination through larger comparative samples and longitudinal empirical analysis.

In conclusion, the Toyota–Honda comparison suggests that adaptation to technological disruption cannot be understood simply as a question of whether firms possess dynamic capabilities. Firms may possess meaningful sensing, seizing, and transforming capabilities while exercising them in fundamentally different ways. The configuration of inherited resources affects how firms interpret change; their interpretation affects how they commit resources; and those commitments subsequently shape their ability to reconfigure again. Dynamic capabilities and resource configuration are therefore closely intertwined in a continuing, recursive process. The capacity to succeed under technological uncertainty may depend not only on the ability to transform resources, but also on the ability to configure them today without unnecessarily constraining tomorrow's strategic choices.

REFERENCES:

- Associated Press, "Honda Canada Postpones Multibillion EV Investment Project in Ontario", 2025, May 13.
- Eisenhardt, K. M., and Martin, J. A., "Dynamic Capabilities: What Are They?", *Strategic Management Journal*, Volume 21, Numbers 10-11, Pages 1105-1121, 2000.
- Financial Times, "Toyota Makes Big Bet on Small Engines in New Hybrid Era", 2024, May 28.
- Helfat, C. E., and Peteraf, M. A., "The Dynamic Resource-Based View: Capability Lifecycles", *Strategic Management Journal*, Volume 24, Number 10, Pages 997-1010, 2003.
- Hoefl, F., "Assessing Dynamic Capabilities of Incumbents in the Face of Unprecedented Industry Transformation: The Case of the Automotive Industry", *Journal of Strategy and Management*, Volume 14, Number 2, Pages 259-283, 2021.
- Honda Motor Co., Ltd., "Summary of 2024 Honda Business Briefing on Direction of Electrification Initiatives and Investment Strategy", 2024, May 16.
- Honda Motor Co., Ltd., "Summary of Honda Briefing on Automobile Electrification Business", 2022, April 12.
- Knobbe, F., and Proff, H., "Dynamic Capabilities in the Automotive Industry Under Digitalization", *International Journal of Automotive Technology and Management*, Volume 20, Number 4, Pages 436-456, 2020.
- Loder, J., Rinscheid, A., and Wüstenhagen, R., "Why Do (Some) German Car Manufacturers Go Electric? The Role of Dynamic Capabilities and Cognitive Frames", *Business Strategy and the Environment*, Volume 33, Number 2, Pages 1129-1143, 2024.
- Ministry of Economy, Trade and Industry, "Future Milestones: Automotive Industry", 2023.
- Reuters, "General Motors, Hyundai to Develop Five Vehicles amid Rising Competition", 2025, August 6.
- Reuters, "Global Carmakers Book \$70 Billion Hit from EV Rollbacks", 2026, February 6.
- Reuters, "GM and Honda Expand Electric Vehicle Partnership", 2022, April 5.
- Reuters, "Japan's Honda to Scale Back on Electric Vehicles, Focus on Hybrids", 2025, May 20.
- Reuters, "Sony, Honda Scale Back EV Venture after Afeela Halt", 2026, April 21.
- Reuters, "Toyota Bets Big on Hybrid-Only Models as EV Demand Slows", 2024, August 15.
- Reuters, "Toyota Opens U.S. Battery Plant, Confirms \$10 Billion Investment Plan", 2025, November 12.
- Reuters, "Toyota Pledges \$2.1 Billion More for EV Battery Plant in North Carolina", 2023, May 31.
- Reuters, "Toyota Showcases Compact Engines Adaptable to Different Fuels in Hybrid Era", 2024, May 28.
- Reuters, "Toyota to Build \$1.29 Billion North Carolina Battery Plant", 2021, December 6.
- Reuters, "Toyota to Outshine Rivals as More Consumers Opt for Hybrids amid EV Slowdown", 2024, February 2.
- Reuters, "Toyota Unveils Sweeping Plans for New Battery Tech, EV Innovation", 2023, June 13.

- Sony Group Corporation, Honda Motor Co., Ltd., and Sony Honda Mobility Inc., “Announcement Regarding the Discontinuation of Development and Launch of Sony Honda Mobility's Electric Vehicle Models and the Review of Its Business Direction”, 2026, March 25.
- Teece, D. J., “China and the Reshaping of the Auto Industry: A Dynamic Capabilities Perspective”, *Management and Organization Review*, Volume 15, Number 1, Pages 177-199, 2019.
- Teece, D. J., “Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance”, *Strategic Management Journal*, Volume 28, Number 13, Pages 1319-1350, 2007.
- Teece, D. J., Pisano, G., and Shuen, A., “Dynamic Capabilities and Strategic Management”, *Strategic Management Journal*, Volume 18, Number 7, Pages 509-533, 1997.
- Toyota Motor Corporation and Idemitsu Kosan Co., Ltd., “Idemitsu and Toyota Announce Beginning of Cooperation toward Mass Production of All-Solid-State Batteries for BEVs”, 2023, October 12.
- Toyota Motor Corporation, “Electrified Technologies—Batteries: Fundamental Technologies to Innovate BEVs”, 2023, June 13.
- Toyota Motor Corporation, “Media Briefing on Battery EV Strategies”, 2021, December 14.

AUTHOR PROFILE:

Dr. Woonghee Lee earned his Ph.D. at the Ohio State University, Columbus in 1998. Currently he is a professor of strategic management at Hanyang University, Seoul, Korea.

DIGITAL SERVQUAL FRAMEWORK FOR EXAMINING CUSTOMER PERCEPTIONS OF DIGITAL SERVICES IN ONLINE ACCOMMODATION PLATFORMS

Susita Carroll, Winston-Salem State University, North Carolina, U.S.A.

dx.doi.org/10.18374/JABE-26-3.8

ABSTRACT

Online accommodation platforms have become increasingly popular and have changed how people experience lodging reservation services. These platforms are not only used for online booking but also provide digital features such as online payments, digital check-in, instant communication with hosts, and online review systems. As the use of digital check-in continues to grow, understanding how customers perceive this feature has become increasingly important. Despite its growing adoption, limited research has specifically examined customer perceptions of digital check-in services within online accommodation platforms. This paper addresses this gap by proposing a Digital SERVQUAL (D-SERVQUAL) framework that extends the traditional SERVQUAL model by integrating technology-enabled service quality dimensions relevant to digital check-in. The proposed framework may also help accommodation providers better understand customer perceptions of digital check-in when evaluating investments in digital services and their potential value to the organization.

Keywords: Digital Services, hospitality, quality, online accommodation

1. INTRODUCTION

The rapid development of digital services in the hospitality industry has transformed service delivery (Dominguez et al., 2026; Lee, 2016; Lema, 2009). This transformation has enabled accommodation platforms to evolve beyond reservation services to support multiple stages of customer experience. In addition to reservation services, these platforms now provide technology-enabled functions such as online payments, digital check-in and check-out, instant online chatbox, and online review systems (Askenazy & Bourgeois, 2026; Buhalis & Leung, 2018; Dominguez et al., 2026; Meuter et al., 2000). These digital features have become increasingly important in shaping customer experience and repurchase intentions among users of online accommodation platforms (Gummerus et al., 2019; Liang et al., 2018; Mao & Lyu, 2017).

Among the changes in technology-related services, digital check-in represents a significant shift in service delivery by transforming the traditional in-person check-in process into a digital service experience. Digital check-in reduces dependence on direct interactions with service personnel and may contribute to operational efficiency. Although digital check-in is expected to improve customer convenience, little is known about how customers perceive these changes in service delivery. This shift raises important questions about how customers evaluate service quality in digital hospitality environments.

Although the SERVQUAL model has been extensively applied to evaluate service quality within the hospitality industry, its original dimensions were developed in the context of traditional face-to-face service encounters. These dimensions may not fully capture customers' evaluations of technology-enabled service experiences. Similarly, while the Technology Acceptance Model (TAM) explains users' acceptance of digital technologies through perceived usefulness and perceived ease of use, it does not directly assess service quality from a hospitality perspective.

Research on online accommodation platforms has primarily focused on overall technology acceptance, customer satisfaction, and repurchase intentions (Liang et al., 2018; Mao & Lyu, 2017; Minh et al., 2015; Venkatesh et al., 2012; Meuter et al., 2000; Tang et al., 2026). Less attention has been given to customer perceptions of digital check-in as a distinct service encounter within these platforms. Digital check-in represents more than a technological function; it is a service touchpoint that influences customers'

evaluations of convenience, efficiency, reliability, and trust. Therefore, there is a need for a conceptual framework that integrates both service quality and technology acceptance perspectives to better understand customer perceptions of digital check-in.

This paper proposes that understanding customer perceptions of digital check-in requires the integration of both service quality and technology acceptance perspectives. The study adapts the traditional SERVQUAL framework into a Digital SERVQUAL (D-SERVQUAL) model that incorporates technology-related service characteristics. This study provides a conceptual framework for examining how customers perceive digital check-in and how these perceptions contribute to service quality and customer satisfaction in online accommodation platforms.

2. THEORETICAL PARADIGM AND LITERATURE REVIEW

This section reviews the theoretical foundations and relevant literature that underpin the proposed Digital SERVQUAL (D-SERVQUAL) framework for examining customer perceptions of digital check-in in online accommodation platforms. It includes perspectives from the SERVQUAL model, the Technology Acceptance Model (TAM), and Self-Service Technology (SST) theory to explain how customers evaluate technology-enabled service encounters. The section also synthesizes existing research on digital hospitality services and identifies the research gap that justifies the development of a Digital SERVQUAL framework tailored to digital check-in.

The SERVQUAL model is one of the most widely applied frameworks for evaluating service quality in hospitality research. Developed by Parasuraman et al. (1988, 1994), the model conceptualizes service quality through five dimensions: reliability, responsiveness, assurance, empathy, and tangibles. These dimensions evaluate the gap between customers' expectations and their perceptions of actual service performance and have been extensively used to assess customer satisfaction in traditional hospitality settings. The model assumes that customers evaluate service quality by comparing expected service with the service they actually receive. Although SERVQUAL has been widely adopted in hospitality research, it was developed primarily for face-to-face service encounters and may not fully capture customers' evaluations of technology-enabled services within digital hospitality environments.

To address this limitation, the Technology Acceptance Model (TAM) provides a complementary perspective for understanding customers' acceptance of digital technologies. Davis (1989) proposed that technology adoption is primarily determined by two constructs: perceived usefulness and perceived ease of use. Perceived usefulness refers to the extent to which customers believe that technology improves the efficiency and effectiveness of service delivery, whereas perceived ease of use reflects the degree to which the technology is perceived as simple and effortless to use. These constructs have been applied to explain customers' acceptance of digital hospitality technologies, including online booking systems, mobile applications, and smart hotel services (Buhalis & Law, 2008; Tang et al., 2026; Venkatesh & Davis, 2000; Venkatesh et al., 2003).

Self-Service Technology (SST) provides an additional theoretical perspective for understanding digital services. Meuter et al. (2000) define SSTs as technological interfaces that enable customers to perform services independently without direct employee assistance. Within online accommodation platforms, digital check-in represents a self-service technology that allows guests to complete the arrival process remotely through digital platforms, reducing dependence on front desk personnel while increasing convenience and operational efficiency. Previous studies have shown that customers evaluate SSTs based on attributes such as reliability, convenience, security, and ease of use, which influence customer satisfaction and continued usage intentions (Meuter et al., 2000; Oh et al., 2013; Tarnanidis et al., 2026).

These findings suggest that customers' evaluations of digital services are influenced not only by traditional service quality but also by the quality of technology-mediated interactions.

Collectively, SERVQUAL, TAM, and SST provide a comprehensive theoretical foundation for examining customer perceptions of digital check-in. SERVQUAL explains how customers evaluate service quality,

TAM explains customers' acceptance of digital technologies, and SST explains customers' evaluations of self-service technologies. Integrating these perspectives provides the theoretical basis for developing a Digital SERVQUAL (D-SERVQUAL) framework that reflects the characteristics of digital check-in within online accommodation platforms.

Previous studies have applied the original SERVQUAL model to evaluate service quality in the hospitality industry and have consistently demonstrated its effectiveness in measuring customer satisfaction. Research has shown that the reliability, responsiveness, assurance, empathy, and tangibles dimensions positively influence customers' perceptions of service quality, satisfaction, and behavioural intentions, including customer loyalty and revisit intention (Dominguez et al., 2026; Minh et al., 2015; Parasuraman et al., 1988, 1994). These findings demonstrate that SERVQUAL remains an effective framework for evaluating service quality in conventional hospitality settings.

As hospitality services have become increasingly digitalized, researchers have extended the application of service quality theories to technology-enabled service encounters. Studies applying the Technology Acceptance Model (TAM) have consistently found that perceived usefulness and perceived ease of use significantly influence customers' acceptance of digital hospitality technologies, including online booking systems, mobile applications, and smart hotel services (Davis, 1989; Dominguez et al., 2026; Venkatesh & Davis, 2000; Venkatesh et al., 2003). These findings suggest that customers' evaluations of digital services depend not only on service quality but also on how effectively the technology supports the service experience.

Similarly, research on Self-Service Technologies (SSTs) has demonstrated that customers evaluate technology-enabled services based on convenience, reliability, security, ease of use, and perceived control (Meuter et al., 2000; Oh et al., 2013). Within hospitality, digital check-in enables guests to complete the arrival process independently, providing greater flexibility while reducing waiting times and unnecessary face-to-face interactions. Customers' perceptions of digital check-in are shaped by both the quality of the technology and the quality of the service delivered through that technology (Gummerus et al., 2019; Lee et al., 2025; Rosenbaum & Wong, 2015; Shiwen et al., 2022; Tarnanidis et al., 2026).

Existing research has primarily focused on technology adoption, online booking systems, or overall hospitality service quality rather than customer perceptions of digital check-in as a distinct service encounter within online accommodation platforms. As digital check-in becomes an increasingly important component of the guest experience, there is a need for a framework that integrates traditional service quality dimensions with technology-enabled service characteristics. Therefore, this paper develops a Digital SERVQUAL (D-SERVQUAL) framework to examine customer perceptions of digital check-in within online accommodation platforms.

3. PROPOSED SERVQUAL (D-SERVQUAL) MODEL

Building upon the theoretical foundations of SERVQUAL, the Technology Acceptance Model (TAM), and Self-Service Technology (SST) theory, this paper proposes a Digital SERVQUAL (D-SERVQUAL) framework for examining customer perceptions of digital check-in within online accommodation platforms. The framework extends the traditional SERVQUAL model by incorporating technology-enabled service quality dimensions that reflect the distinctive characteristics of digital service delivery. While the original SERVQUAL dimensions remain relevant for evaluating service performance, they do not fully capture customers' evaluations of technology-mediated service encounters in which digital systems perform functions traditionally delivered by frontline employees.

The proposed D-SERVQUAL framework integrates both service quality and technology acceptance perspectives to provide a more comprehensive evaluation of digital check-in. The framework recognizes that customers evaluate digital check-in not only according to the quality of the service provided but also according to the functionality, usability, and security of the digital platform through which the service is delivered. Customer perceptions are shaped by both traditional service quality attributes and technology-related characteristics that influence the overall digital service experience.

The framework comprises six interrelated dimensions: reliability, responsiveness, security, perceived usefulness, perceived ease of use, and autonomy. Reliability reflects the ability of the digital check-in system to process reservations accurately and consistently without operational errors. Responsiveness refers to the efficiency and timeliness with which the system confirms reservations and completes the check-in process. Security captures customers' confidence that their personal information, payment details, and reservation data are adequately protected throughout the digital transaction. Perceived usefulness reflects the extent to which customers believe that digital check-in enhances the efficiency, convenience, and effectiveness of the arrival process, whereas perceived ease of use represents the degree to which the system is simple, intuitive, and requires minimal effort to operate. Finally, autonomy refers to customers' ability to independently control and manage the check-in process according to their own preferences, without requiring assistance from hotel employees.

Unlike the original SERVQUAL model, which was developed to evaluate face-to-face service encounters, the proposed D-SERVQUAL framework acknowledges that digital hospitality services require customers to simultaneously evaluate service quality and technology performance. As digital check-in increasingly becomes a standard feature of online accommodation platforms, customers' perceptions are influenced by the interaction between these two perspectives. The integration of service quality and technology-related dimensions provides a more appropriate conceptual framework for evaluating digital check-in within contemporary hospitality environments.

Figure 1 presents the proposed Digital SERVQUAL (D-SERVQUAL) framework.

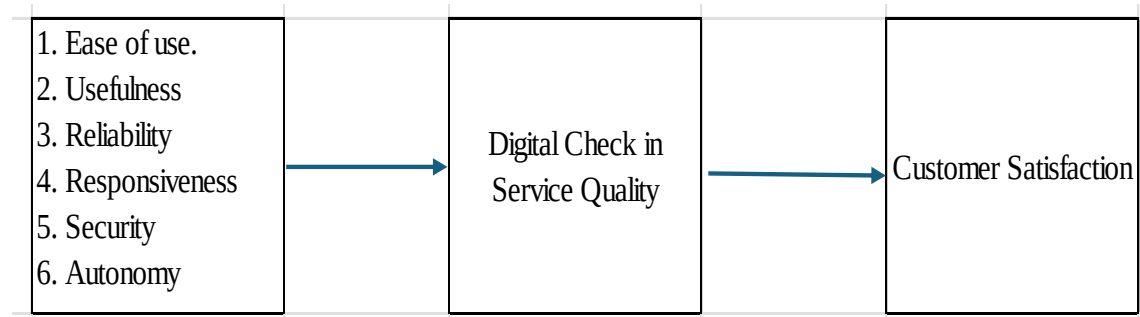


Figure 1: Proposed Model

Table 1 presents the conceptual definitions and sub-dimensions of each construct. Together, the framework provides a theoretical foundation for understanding customer perceptions of digital check-in and offers a basis for future empirical research examining digital service quality in online accommodation platforms.

Constructs and Dimensions	Definition	Key References
Self-Service Digital Check-In Quality	Refers to customers' overall evaluation of a digital hotel service that enables them to complete the check-in process independently before or upon arrival without relying on front desk staff.	Buhalis & Leung (2018); Meuter et al. (2000); Tarnanidis et al. (2026)
Perceived Usefulness	Refers to the extent to which customers believe that using the digital check-in system enhances the efficiency, convenience, and overall effectiveness of the hotel arrival process.	Davis (1989); Dominguez et al. (2026); Venkatesh & Davis (2000); Venkatesh et al. (2003)
Perceived Ease of Use	Refers to the degree to which customers perceive the digital check-in system as easy to understand, straightforward to operate, and requiring minimal effort.	Davis (1989); Dominguez et al. (2026); Venkatesh & Davis (2000); Venkatesh et al. (2003)
Reliability	Refers to the ability of the digital check-in system to perform accurately, consistently, and dependably while processing reservations and completing the check-in procedure without errors.	Parasuraman et al. (1988); Zeithaml et al. (2002); Parasuraman et al. (2005); Blut (2016)
Responsiveness	Refers to the ability of the digital check-in system to provide prompt processing, immediate confirmation, real-time feedback, and timely assistance throughout the digital check-in process.	Blut (2016); Parasuraman et al. (1988); Parasuraman et al. (2005); Zeithaml et al. (2002);
Security	Refers to the extent to which customers believe that the digital check-in system adequately protects their personal information, payment details, and reservation data throughout the digital check-in process.	Parasuraman et al. (2005); Blut (2016); Venkatesh et al. (2012)
Autonomy	Refers to customers' perception that the digital check-in system provides the freedom and independence to manage the hotel arrival process and make check-in decisions without assistance from hotel staff.	Dabholkar (1996); Deci & Ryan (2000); Meuter et al. (2000);
Perceived Satisfaction (<i>Outcome Construct</i>)	Refers to customers' overall evaluation of their digital check-in experience based on the extent to which the service meets or exceeds their expectations.	Fornell et al. (1996); Oliver (1980); Tarnanidis et al. (2026)

Table 1: Construct definitions.

4. ANALYSIS AND DISCUSSION

This paper proposes a Digital SERVQUAL (D-SERVQUAL) framework for examining customer perceptions of digital check-in in online accommodation platforms by integrating constructs from the SERVQUAL model, the Technology Acceptance Model (TAM), and Self-Service Technology (SST) theory. The proposed model incorporates reliability, responsiveness, security, perceived usefulness, perceived ease of use, and autonomy as key dimensions of digital check-in service quality.

The integration of these dimensions reflects the changing nature of service delivery in hospitality. In a traditional service encounter, customers often evaluate service quality based on their interactions with employees. In a digital check-in environment, however, part of this interaction is replaced by technology. As a result, customers may evaluate the service not only based on whether it is reliable and responsive, but also on whether the technology is useful, easy to use, secure, and allows them to complete the process independently. The proposed framework therefore recognizes that digital check-in involves both a service experience and a technology experience.

The proposed dimensions may also work together in shaping customers' overall perceptions of digital check-in. For example, a system may be easy to use but may still result in a poor service experience if it is unreliable or does not provide timely confirmation.

Similarly, customers may find digital check-in convenient but may be reluctant to use it if they have concerns about the security of their personal or payment information. Autonomy may further influence the experience by giving customers greater control over when and how they complete the check-in process. These considerations suggest that evaluating digital check-in through a single service or technology dimension may not fully explain how customers perceive the overall experience.

Following the collection of empirical data, the proposed D-SERVQUAL framework will be empirically analyzed to examine the relationships among the proposed dimensions, Digital Check-In Service Quality, and perceived satisfaction. The analysis will assess the reliability and validity of the constructs and test the proposed relationships within the conceptual framework. The empirical findings will provide evidence regarding the applicability of the proposed D-SERVQUAL framework in the context of digital check-in services.

The empirical analysis will also help determine which dimensions have stronger relationships with customers' evaluations of digital check-in service quality and perceived satisfaction. This is important because customers may not place equal importance on every aspect of the digital experience. Identifying these relationships can provide a clearer understanding of how customers evaluate digital check-in and whether the proposed dimensions adequately represent service quality in a technology-enabled hospitality environment.

The proposed framework contributes to the hospitality literature in several ways. First, it extends the traditional SERVQUAL model by adapting service quality dimensions to the context of digital hospitality services. Second, it integrates technology acceptance and self-service perspectives into a single conceptual framework, providing a more comprehensive understanding of customer evaluations of digital check-in. Third, it conceptualizes Digital Check-In Service Quality (D-SERVQUAL) as a construct that influences customers' perceived satisfaction, offering a theoretical foundation for future empirical studies.

More broadly, the framework responds to the increasing movement from employee-assisted service encounters toward technology-enabled and self-service experiences in hospitality. As digital check-in becomes more common, understanding service quality requires attention to both the service provided and the digital system through which customers receive that service.

The D-SERVQUAL framework provides a basis for examining these two aspects together and may help explain how digital service encounters influence customers' overall evaluations of their hospitality experience.

5. MANAGERIAL IMPLICATIONS

The growing adoption of digital check-in provides important opportunities for accommodation providers to improve the overall guest experience and operational efficiency. From a managerial perspective, the framework provides guidance for accommodation providers and online platform operators seeking to improve digital check-in services. By understanding the dimensions that shape customers' evaluations, managers can prioritize improvements in system reliability, responsiveness, security, usability, and customer autonomy, thereby enhancing the overall digital guest experience and customer satisfaction.

Service providers who decide to incorporate these digital features can strengthen their competitiveness by offering more convenient and flexible check-in services. At the same time, the application can improve operational efficiency by reducing unnecessary costs, minimizing routine front-desk activities, and allowing employees to focus on more complex tasks. Together, these improvements can enhance service performance, improve customer satisfaction, and increase the competitiveness of accommodation providers.

The effective implementation of digital check-in can create value for both customers and accommodation providers. Service quality can be improved, and customers can have a more convenient and efficient experience (Shampa & Hossain, 2026; Tang et al., 2026; Tarnanidis et al., 2026). Digital check-in can also provide customers with greater flexibility and control over the arrival process while reducing the time required for traditional front-desk interactions. For accommodation providers, the effective use of digital check-in can support more efficient service delivery by reducing routine administrative activities and allowing employees to focus on other important aspects of the guest experience. As customer expectations for convenient and technology-enabled services continue to develop, providing reliable and easy-to-use digital check-in options may also help accommodation providers remain competitive. Therefore, the successful implementation of digital check-in has the potential to improve both the customer experience and the operational performance of accommodation providers.

6. LIMITATIONS AND FUTURE RESEARCH

This study has several limitations. First, the proposed D-SERVQUAL framework is conceptual and has not been empirically validated. The proposed relationships among the dimensions, Digital Check-In Service Quality, and perceived satisfaction remain theoretical and require empirical examination.

Second, the framework is developed for digital check-in within online accommodation platforms. Its applicability to other hospitality technologies, such as digital check-out, mobile concierge services, or smart room technologies, has not been examined and may require further adaptation.

Third, although the framework integrates constructs from SERVQUAL, TAM, and SST theory, other factors such as trust, perceived risk, digital literacy, customer experience, or cultural differences may also influence customers' perceptions of digital check-in and were not incorporated into the present framework.

Future research should empirically validate the proposed D-SERVQUAL framework using quantitative research methods such as confirmatory factor analysis (CFA) and structural equation modelling (SEM). Such studies would enable researchers to examine the reliability and validity of the proposed dimensions and assess the relationships between Digital Check-In Service Quality and perceived satisfaction.

Empirical testing could also provide a better understanding of which dimensions have the strongest influence on customers' evaluations of digital check-in services.

Future studies may also investigate the applicability of the framework across different hospitality contexts, including hotels, resorts, serviced apartments, and peer-to-peer accommodation platforms, to determine whether customer perceptions vary across accommodation types. Comparative studies across countries and cultural settings would enhance understanding of how cultural and technological differences influence customer evaluations of digital check-in. This may also help determine whether the framework can be applied consistently across different types of accommodation services and customer populations.

In addition, future research could extend the framework by incorporating other technology-related constructs, such as trust, perceived risk, perceived value, digital competence, or customer experience. Future studies could also test the framework across different locations, cultures, and age groups to better understand whether customers perceive digital check-in differently. For example, younger and older customers may have different levels of comfort and experience with digital services. Cultural differences may also shape customers' expectations and preferences when using digital check-in.

Examining these differences could help determine whether the proposed D-SERVQUAL framework can be applied across different customer groups and hospitality settings. These comparisons could provide additional insight into how individual and contextual differences shape perceptions of digital service quality. They may also help identify whether certain dimensions of the framework are more important for specific customer groups or hospitality settings. This would provide a broader understanding of how customers evaluate digital check-in services across different service environments.

REFERENCES:

- Askenazy, P., & Bourgeois, A. (2026). Accommodation via peer-to-peer platforms in national accounts. *Annals of Tourism Research Empirical Insights*, 7(1), 100207.
- Blut, M., "E-service Quality: Development of a Hierarchical Model", *Journal of Retailing*, Volume 92, Number 4, Pages 500–517, 2016.
- Buhalis, D. and Law, R., "Progress in Information Technology and Tourism Management: 20 Years On and 10 Years After the Internet—The State of eTourism Research", *Tourism Management*, Volume 29, Number 4, Pages 609–623, 2008.
- Buhalis, D. and Leung, R., "Smart Hospitality—Interconnectivity and Interoperability Towards an Ecosystem", *International Journal of Hospitality Management*, Volume 71, Pages 41–50, 2018.
- Dabholkar, P. A., "Consumer Evaluations of New Technology-Based Self-Service Options: An Investigation of Alternative Models of Service Quality", *International Journal of Research in Marketing*, Volume 13, Number 1, Pages 29–51, 1996.
- Davis, F. D., "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology", *MIS Quarterly*, Volume 13, Number 3, Pages 319–340, 1989.
- Deci, E. L. and Ryan, R. M., "The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self-Determination of Behavior", *Psychological Inquiry*, Volume 11, Number 4, Pages 227–268, 2000.
- Dominguez, J. M. G., Manalon, B. L. A., Martinez, T. J. D., Pablo, S. C. D., Viray, A. D., & Lunar, R. A. (2026). System quality of digitalized bookings in four-star hotels and its impact on guests' acceptance. *Journal of Hospitality, Tourism & Cultural Research*, 2(2), 68–90.
- Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J. and Bryant, B. E., "The American Customer Satisfaction Index: Nature, Purpose, and Findings", *Journal of Marketing*, Volume 60, Number 4, Pages 7–18, 1996.
- Gummerus, J., Lipkin, M., Dube, A. and Heinonen, K., "Technology in Use: Characterizing Customer Self-Service Devices (SSDS)", *Journal of Services Marketing*, Volume 33, Number 1, Pages 44–56, 2019.
- Lee, L. Y. S., "Hospitality Industry Web-Based Self-Service Technology Adoption Model: A Cross-Cultural Perspective", *Journal of Hospitality & Tourism Research*, Volume 40, Number 2, Pages 162–197, 2016.

- Lee, S., Shulga, L. V., Busser, J. A. and Hwang, J., "Self-Service Hotel Check-In: Is There Still a Role for the In-Person Experience?", *International Journal of Hospitality & Tourism Administration*, Volume 26, Number 5, Pages 936–963, 2025.
- Lema, J. D., "Preparing Hospitality Organizations for Self-Service Technology", *Journal of Human Resources in Hospitality & Tourism*, Volume 8, Number 2, Pages 153–169, 2009.
- Liang, L. J., Choi, H. C. and Joppe, M., "Exploring the Relationship Between Satisfaction, Trust and Switching Intention, Repurchase Intention in the Context of Airbnb", *International Journal of Hospitality Management*, Volume 69, Pages 41–48, 2018.
- Meuter, M. L., Ostrom, A. L., Roundtree, R. I. and Bitner, M. J., "Self-Service Technologies: Understanding Customer Satisfaction with Technology-Based Service Encounters", *Journal of Marketing*, Volume 64, Number 3, Pages 50–64, 2000.
- Minh, N. V., Ha, N. T., Anh, P. C. and Matsui, Y., "Service Quality and Customer Satisfaction: A Case Study of Hotel Industry in Vietnam", *Asian Social Science*, Volume 11, Number 10, Pages 73–85, 2015.
- Mao, Z. and Lyu, J., "Why Travelers Use Airbnb Again? An Integrative Approach to Understanding Travelers' Repurchase Intention", *International Journal of Contemporary Hospitality Management*, Volume 29, Number 9, Pages 2464–2482, 2017.
- Oh, H., Jeong, M. and Baloglu, S., "Tourists' Adoption of Self-Service Technologies at Resort Hotels", *Journal of Business Research*, Volume 66, Number 6, Pages 692–699, 2013.
- Oliver, R. L., "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions", *Journal of Marketing Research*, Volume 17, Number 4, Pages 460–469, 1980.
- Parasuraman, A., Zeithaml, V. A. and Berry, L. L., "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality", *Journal of Retailing*, Volume 64, Number 1, Pages 12–40, 1988.
- Parasuraman, A., Zeithaml, V. A. and Berry, L. L., "Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research", *Journal of Marketing*, Volume 58, Number 1, Pages 111–124, 1994.
- Parasuraman, A., Zeithaml, V. A. and Malhotra, A., "E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality", *Journal of Service Research*, Volume 7, Number 3, Pages 213–233, 2005.
- Rosenbaum, M. S. and Wong, I. A., "If You Install It, Will They Use It? Understanding Why Hospitality Customers Take 'Technological Pauses' from Self-Service Technology", *Journal of Business Research*, Volume 68, Number 9, Pages 1862–1868, 2015.
- Shampa, M. T. J., & Hossain, M. S. (2026). Analyzing factors influencing smart technology adoption in an emerging economy's tourism and hospitality sector. *Telematics and Informatics Reports*, 21, 100308.
- Shiwen, L., Kwon, J. and Ahn, J., "Self-Service Technology in the Hospitality and Tourism Settings: A Critical Review of the Literature", *Journal of Hospitality & Tourism Research*, Volume 46, Number 6, Pages 1220–1236, 2022.
- Tang, T., Lee, J., & Chung, N. (2026). Beyond ease of use: How smart hotel technology amenities create memorable experiences and revisit intention. *The Journal of Internet Electronic Commerce Research*, 26(2), 327–351

- Tarnanidis, T., Ismyrlis, V., Manda, V. K., & Sousa, B. (2026). Redefining service quality in the digital experience era. *Journal of Hospitality and Tourism Horizons*, 2(1), 23–42
- Venkatesh, V. and Davis, F. D., “A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies”, *Management Science*, Volume 46, Number 2, Pages 186–204, 2000.
- Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D., “User Acceptance of Information Technology: Toward a Unified View”, *MIS Quarterly*, Volume 27, Number 3, Pages 425–478, 2003.
- Venkatesh, V., Thong, J. Y. L. and Xu, X., “Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology”, *MIS Quarterly*, Volume 36, Number 1, Pages 157–178, 2012.
- Zeithaml, V. A., Parasuraman, A. and Malhotra, A., “Service Quality Delivery Through Websites: A Critical Review of Extant Knowledge”, *Journal of the Academy of Marketing Science*, Volume 30, Number 4, Pages 362–375, 2002.

AUTHOR PROFILE:

Susita Carroll is Professor of Management in the Department of Business Administration at Winston-Salem State University, North Carolina, USA. She earned her Ph.D. from the University of Toledo, Ohio. Her research interests include supply chain management, hospitality management, service quality, digital transformation, and technology adoption. Her scholarly work has been published in peer-reviewed journals, including *International Journal of Contemporary Hospitality Management*, *International Journal of Integrated Supply Management*, *Journal of Managerial Issues*, and *Journal of Case Studies*.