STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, August 2026, pp 1219–1242.
Published online in International Academic Press (www.IAPress.org)

Determinants of life insurance market share: a comprehensive panel econometric and machine learning analysis of seven industrialized countries (2005-2022)

Mohamed Ibraheem [1], Hebatallah Hassan [2], Sayed Youssef [3,*], Mousa Gowfal Selmey [4], Noura M. Saad [5]

[1] *Department of Applied Statistics and Insurance, Faculty of Commerce, Mansoura University, Egypt*
[2] *Department of Economics and Finance, College of Business Administration, Taif University, Saudi Arabia*
[3] *College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia*
[4] *Department of Business Administration, College of Business, Jouf University, Sakaka, Saudi Arabia*
[5] *5Department of Statistics, Mathematics and Insurance, Faculty of Commerce, Kafrelsheikh University, Kafrelsheikh 33516, Egypt*

Abstract This study analyzes the determinants affecting life insurance market share in seven prominent industrialized countries: France, Germany, Italy, Japan, Spain, Sweden, and the United Kingdom from 2005 to 2022. We employ a comprehensive panel dataset with 126 country-year observations and apply advanced methodological approaches, including two-way fixed effects, system GMM dynamic panel estimation, panel cointegration, Granger causality tests, threshold regression, and ensemble machine learning (XGBoost) with SHAP interpretability analysis. The empirical results reveal six main findings. First, retention ratios exhibit a strong positive and statistically significant association with life insurance market share (coefficient = 1.512, $p < 0.01$), with system GMM estimates confirming this relationship. Second, reinsurance acceptance demonstrates a strong negative association (coefficient = -7.845, $p < 0.01$), reflecting a fundamental trade-off between primary insurance and reinsurance functions. Third, our novel Reinsurance Dependency Index captures nonlinear interaction effects (coefficient = 589.46, $p < 0.05$). Fourth, life insurance markets display extraordinary persistence ($\gamma = 0.876$), with approximately 88% of the previous year's market share carrying over to the current year. Fifth, threshold analysis identifies a critical retention threshold of 88.2%, above which retention effects strengthen substantially. Sixth, the COVID-19 pandemic caused a modest immediate decline of 0.89 percentage points, with high-retention markets demonstrating greater resilience. While our empirical strategy addresses several sources of endogeneity through system GMM and instrumental variables estimation, we acknowledge that the results are most appropriately interpreted as associations consistent with causal relationships rather than definitive causal identification. The findings have important implications for insurance executives, regulators, and policymakers seeking to understand the strategic determinants of life insurance market development.

Keywords Life Insurance, Retention Ratio, Reinsurance Acceptance, Panel Data, System GMM, Machine Learning, Threshold Regression, COVID-19, Cross-Country Analysis, Actuarial Risk

AMS 2010 subject classifications 62P05, 62M10, 62J05, 91B30

DOI: 10.19139/soic-2310-5070-4020

1. Introduction

1.1. Background and Motivation

The global insurance industry serves as a cornerstone of modern financial systems, with life insurance playing a critical role in long-term savings, mortality protection, and intergenerational wealth transfer. According to the

*Correspondence to: Sayed Youssef (Email: smyahmed@imamu.edu.sa). Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Swiss Re Institute, global life insurance premiums reached approximately USD 3.1 trillion in 2023, representing nearly 55% of total global insurance premiums [1, 2, 3]. However, the relative importance of life insurance varies dramatically across countries due to differences in regulatory frameworks, pension system designs, cultural preferences, and industry structure. Understanding the determinants of life insurance market share is not merely an academic exercise; it has profound practical implications. Insurance executives can make better decisions about product mix, underwriting policies, and reinsurance purchasing when they understand which strategic variables affect market share. For regulators, understanding how retention and reinsurance decisions affect market structure can inform capital requirements and systemic risk monitoring. For investors, changes in market share signal growth opportunities and competitive positioning [4, 5, 6, 7].

The existing literature contains a significant gap that our research addresses. While extensive research has examined macroeconomic determinants of life insurance demand, such as GDP per capita, interest rates, inflation, and demographic structure, industry-specific characteristics have received comparatively little attention. Specifically, how do retention ratios (the proportion of premiums retained by insurers rather than ceded to reinsurers) and reinsurance acceptance affect life insurance market share?

Furthermore, the COVID-19 pandemic provided a distinct exogenous shock for studying the resilience of different insurance market structures [8, 9, 10]. Unlike the 2008-2009 financial crisis, which originated in financial sector vulnerabilities, the pandemic was a genuine external shock. This study provides the first systematic comparison of how markets with different retention characteristics responded to these two distinct crises.

1.2. Research Questions and Objectives

This study is guided by four primary research questions:

RQ1: What is the empirical relationship between retention ratios and life insurance market share across developed insurance markets?

RQ2: How do outward and inward reinsurance activities affect the relative size of life insurance sectors?

RQ3: Can a composite measure of reinsurance dependency, capturing the interaction between retention and reinsurance, provide superior explanatory power for cross-country variation in life insurance share?

RQ4: How did the COVID-19 pandemic affect life insurance market share, and did high-retention markets demonstrate greater resilience?

To address these questions, we construct a balanced panel dataset spanning 18 years (2005-2022) for seven countries that collectively represent approximately 65% of global life insurance premiums (Swiss Re Institute, 2023). We employ an advanced methodological approach, including two-way fixed effects panel regression, system GMM dynamic panel estimation, panel cointegration, Granger causality tests, threshold regression, ensemble machine learning with SHAP interpretability analysis, and event study analysis of the pandemic period.

1.3. Contributions to the Literature

This paper makes six distinct contributions to the insurance economics literature.

First, we introduce and validate a novel metric, the Reinsurance Dependency Index (RDI), defined as the ratio of reinsurance accepted to retention. This index captures the strategic orientation of an insurance market toward either retaining risk (low RDI) or functioning as a reinsurance hub (high RDI). We acknowledge that interpretation of RDI should be contextualized with knowledge of its components, as a high RDI may result from either high reinsurance acceptance or low retention.

Second, we provide rigorous econometric evidence on the relationship between operational strategies (retention and reinsurance) and market outcomes (life insurance share) using system GMM estimation that addresses endogeneity concerns. While our empirical strategy addresses several sources of endogeneity, we interpret our findings as evidence consistent with causal relationships rather than definitive causal identification.

Third, we identify a critical retention threshold (88.2%) using Hansen's panel threshold regression, revealing nonlinear effects where retention strategies yield increasing returns above this threshold.

Fourth, we document extraordinary persistence in life insurance markets ($\gamma = 0.876$), demonstrating that supply-side factors create longer-lasting stickiness than previously recognized.

Fifth, we develop an ensemble machine learning model (XGBoost + fixed effects) that achieves superior predictive performance ($R^2 = 0.901$, MAPE = 6.1%), substantially outperforming traditional econometric approaches. We complement this with SHAP (SHapley Additive exPlanations) analysis to enhance interpretability and demonstrate the direction of effects.

Sixth, we provide the first systematic comparison of pandemic versus financial crisis effects on life insurance markets, demonstrating that high-retention markets showed greater resilience during COVID-19.

1.4. Paper Structure

The remainder of this paper is organized as follows. Section 2 reviews the theoretical and empirical literature. Section 3 describes our data sources and variable construction. Section 4 presents descriptive statistics and preliminary analysis. Section 5 details our enhanced econometric methodology. Section 6 reports comprehensive empirical results. Section 7 presents machine learning enhancement with SHAP interpretability analysis. Section 8 provides a COVID-19 impact analysis. Section 9 discusses implications. Section 10 concludes.

2. LITERATURE REVIEW

2.1. Theoretical Foundations of Life Insurance Demand

The theoretical foundations of life insurance originate from the groundbreaking research of Yaari (1965), who demonstrated that life insurance purchases enable consumption smoothing over uncertain lifetimes [6]. Fischer (1973) and Campbell (1980) subsequently incorporated bequest motives and tax considerations [11, 12, 13]. The standard life-cycle model predicts that life insurance demand increases with income, number of dependents, and bequest motives, while decreasing with wealth and availability of alternative savings vehicles.

However, the theoretical literature has paid limited attention to supply-side factors. Market structure including the extent of risk retention versus reinsurance affects the price, availability, and variety of life insurance products. In markets where insurers retain most risks (high retention), they have stronger incentives to price accurately and manage underwriting quality through three primary mechanisms: (1) incentive alignment between underwriting decisions and profitability, (2) capital accumulation through retained earnings that enables market expansion, and (3) reputation and market positioning effects where insurers with demonstrated risk management capabilities attract more policyholders. Conversely, markets with extensive reinsurance may face moral hazard problems, as ceding insurers have reduced incentives to screen and monitor risks [14].

The theoretical mechanisms linking retention to market share can be formalized as follows. First, higher retention creates stronger incentives for accurate underwriting and effective risk management, as insurers bear a larger proportion of the consequences of their underwriting decisions. Second, retained premiums build internal capital over time, enabling insurers to expand their product offerings and market presence. Third, high-retention insurers may develop superior risk assessment capabilities and reputational advantages that attract policyholders. These mechanisms suggest a positive relationship between retention and market share, but the relationship may be nonlinear if fixed costs or threshold effects are present.

2.2. Empirical Determinants of Life Insurance Consumption

A substantial empirical literature examines macroeconomic and demographic determinants of life insurance consumption. Outreville (2013) provides a comprehensive meta-analysis of 85 empirical papers, identifying GDP per capita, financial development, inflation, and the dependency ratio as the most consistent predictors [15].

Beck and Webb (2003), using panel data from 68 countries over 1961-2000, find that life insurance penetration is positively associated with banking sector depth but negatively associated with stock market capitalization, suggesting substitution between insurance and equity investments [16]. Chang and Lee (2012) demonstrate a nonlinear relationship between life insurance development and economic growth, finding that the relationship is positive only beyond a certain threshold of economic development [17]. This threshold effect is consistent with our later findings regarding retention thresholds.

The impact of the post-2008 low-interest-rate environment has received considerable attention. Berdin and Gründl (2015) analyzed European life insurers and found that a 100-basis-point decline in interest rates reduces profitability by 15-20% for traditional products, forcing insurers to shift toward fee-based unit-linked products [18].

2.3. *Retention, Reinsurance, and Market Structure*

The literature on retention ratios and reinsurance is more fragmented, with most studies focusing on determinants of reinsurance purchasing rather than market share consequences. Mayers and Smith (1982) develop a theoretical framework in which reinsurance reduces insolvency risk, provides real services (underwriting, claims handling), and signals managerial quality [19]. Consistent with their predictions, empirical studies find that insurers with higher risk, lower capital, and greater geographic concentration purchase more reinsurance.

However, the relationship between reinsurance and market share has received less attention. Descriptive studies note that Germany's high reinsurance acceptance (approximately 30% of gross premiums) reflects its role as a global reinsurance hub, while Sweden's low reinsurance acceptance (approximately 2%) indicates a primary market focus. This study provides the first systematic econometric analysis of how these strategic choices affect life insurance market share, while recognizing that high reinsurance acceptance may indicate market maturity and specialization rather than weakness in the domestic life insurance sector.

2.4. *Crisis Effects on Insurance Markets*

The 2008-2009 financial crisis and the 2020-2021 COVID-19 pandemic provide two distinct crisis settings for studying insurance market resilience. The financial crisis primarily affected life insurance through wealth effects (declining asset values) and interest rate channels (falling yields on government bonds). Koijen and Yogo (2022) document that variable annuity insurers offering minimum return guarantees were particularly exposed to market-risk pressures during crisis periods [1].

The COVID-19 pandemic introduced novel channels: mortality awareness (potentially increasing demand), income disruption (reducing ability to pay premiums), and operational constraints (lockdowns affecting distribution) [20, 21]. Early evidence suggested heterogeneous effects, with some markets experiencing increased demand for term life insurance while others saw premium declines [22]. This study extends this literature by directly comparing the two crises and examining how market structure (retention characteristics) moderated the impact.

2.5. *Research Gap and Hypotheses*

Despite the rich literature, no previous study has examined the joint effect of retention and reinsurance on life insurance market share using advanced panel econometric methods with causality testing, threshold effects, and machine learning enhancement. Based on theoretical reasoning and descriptive evidence, we formulate four hypotheses:

H1 (Retention Hypothesis): Higher retention ratios are positively associated with life insurance market share. Insurers who retain more risk have stronger incentives to compete effectively and build market presence.

H2 (Reinsurance Hypothesis): Higher reinsurance acceptance is negatively associated with life insurance market share. Markets functioning as reinsurance hubs have relatively smaller domestic life insurance sectors. This reflects a fundamental trade-off between primary insurance and reinsurance functions, and should not be interpreted as implying that reinsurance activities are detrimental.

H3 (Dependency Hypothesis): The Reinsurance Dependency Index, capturing the ratio of reinsurance accepted to retention, has a nonlinear relationship with life insurance market share.

H4 (Resilience Hypothesis): High-retention markets demonstrated greater resilience during the COVID-19 pandemic compared to low-retention markets.

3. DATA AND VARIABLE CONSTRUCTION

3.1. Data Sources and Sample

Data for this study are drawn from official publications of national insurance supervisory authorities for the period 2005-2022:

- **France:** Autorité de Contrôle Prudentiel et de Résolution (ACPR)
- **Germany:** Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin)
- **Italy:** Istituto per la Vigilanza sulle Assicurazioni (IVASS)
- **Japan:** Financial Services Agency (FSA) and The Life Insurance Association of Japan (LIAJ)
- **Spain:** Dirección General de Seguros y Fondos de Pensiones (DGSFP)
- **Sweden:** Finansinspektionen
- **United Kingdom:** Prudential Regulation Authority (PRA) / Bank of England

Our sample period spans 18 years (2005-2022), yielding 126 country-year observations (7 countries $\times$ 18 years). This balanced panel structure is ideal for advanced panel econometric estimation.

The selection of these seven countries was motivated by several considerations. First, they collectively represent approximately 65% of global life insurance premiums (Swiss Re Institute, 2023), ensuring empirical relevance. Second, they exhibit substantial variation in retention ratios (ranging from 84.72% in the UK to 95.97% in Italy) and reinsurance acceptance (from 2.04% in Italy to 24.33% in Germany), providing informative variation for identifying relationships. Third, all are high-income OECD members with comparable regulatory frameworks, facilitating cross-country comparison. Fourth, data availability and quality are consistently high across these countries, with detailed official insurance statistics published annually by supervisory authorities.

Macroeconomic and demographic control variables were sourced from:

- World Bank Development Indicators (GDP per capita, inflation rate, old-age dependency ratio)
- OECD Statistics (long-term government bond yields)
- IMF Financial Development Database (financial development index)

3.2. Variable Definitions

3.2.1. Dependent Variable: Life Insurance Market Share (Life Share) Life Share is defined as the ratio of gross written premiums for life insurance to total gross written premiums (life plus non-life), expressed as a percentage. This measure captures the relative importance of life insurance within each national market.

3.2.2. Independent Variables **Retention Ratio (Retention):** Defined as net written premiums divided by total gross written premiums, expressed as a percentage. A higher retention ratio indicates that insurers retain a larger proportion of premiums after reinsurance cessions.

Reinsurance Acceptance Ratio (Reinsurance): Defined as reinsurance accepted from other insurers divided by total gross written premiums, expressed as a percentage. This measure captures the extent to which a market functions as a reinsurance hub.

Reinsurance Dependency Index (RDI): A novel metric constructed as the ratio of reinsurance accepted to retention (Reinsurance / Retention). Values near zero indicate a primary market that retains risk; values exceeding 0.20 indicate significant reinsurance activity. In our sample, retention ratios consistently remained below 100% (range: 84.72% to 95.97%), ensuring the denominator is always positive. We acknowledge that a high RDI may result from either high reinsurance acceptance (as in Germany) or low retention (as in the UK during certain periods), and interpretation should be contextualized with knowledge of its components.

Relative Penetration Index (RPI): Defined as the natural logarithm of the product of life share and retention ratio: ln(Life Share $\times$ Retention). This metric captures effective market presence after accounting for risk transfer.

3.2.3. Control Variables **Macroeconomic Controls:** To address potential omitted variable bias, we incorporate the following macroeconomic and demographic control variables:

- **GDP per capita (PPP):** Sourced from World Bank Development Indicators, measured in constant international dollars.
- **Long-term government bond yields:** Sourced from OECD and national central banks, expressed as percentages.
- **Inflation rate (CPI):** Sourced from World Bank Development Indicators, measured as annual percentage change.
- **Old-age dependency ratio:** Sourced from World Bank, defined as the ratio of population aged 65+ to working-age population (15-64).
- **Financial development index:** Sourced from IMF Financial Development Database, a composite measure of financial institution and market development.
- **Crisis Indicator (2008-2009):** A binary variable equal to 1 for 2008-2009, 0 otherwise, capturing the global financial crisis.
- **Post-Solvency II Indicator:** A binary variable equal to 1 for 2016 onward for EU countries only, 0 otherwise, capturing the full implementation of Solvency II in European markets.
- **Pandemic Indicator:** A binary variable equal to 1 for 2020-2021, 0 otherwise, capturing the COVID-19 pandemic period. We exclude 2022 from the pandemic period for the main analysis to avoid classification issues, but include it in robustness checks.

3.3. Descriptive Statistics

Table 1 presents summary statistics by country for the full sample period (2005-2022).

Table 1. Summary Statistics by Country (2005-2022)

Country	Life Share (%)	Retention (%)	Reinsurance Accepted (%)	RDI	RPI
France	53.06 (8.41)	88.31 (4.12)	11.46 (3.21)	0.130 (0.041)	8.43 (0.32)
Germany	36.10 (7.23)	87.64 (5.34)	24.33 (6.12)	0.278 (0.087)	8.05 (0.45)
Italy	71.99 (6.87)	95.97 (2.15)	2.04 (0.89)	0.021 (0.009)	8.83 (0.28)
Japan	77.28 (4.56)	90.07 (3.42)	5.72 (1.45)	0.063 (0.018)	8.85 (0.22)
Spain	43.44 (7.89)	90.31 (3.87)	7.00 (2.11)	0.078 (0.025)	8.26 (0.38)
Sweden	65.68 (11.23)	95.26 (2.78)	4.70 (1.34)	0.049 (0.014)	8.73 (0.41)
United Kingdom	72.04 (6.54)	84.72 (5.67)	14.66 (4.23)	0.173 (0.056)	8.69 (0.35)

Note: Standard deviations in parentheses. Source: Authors' calculations based on national supervisory authority data.

Several patterns emerge from Table 1. Japan maintains the highest life share (77.3%), followed by the UK and Italy (approximately 72%). Germany has the lowest life share (36.1%) but the highest RDI (0.278), reflecting its role as a global reinsurance hub. Italy and Sweden show the highest retention ratios (>95%), while the UK shows the lowest retention (84.7%).

3.4. Graphical Analysis

Figure 1: Historical Trends of Reinsurance Acceptance Rates by Country (2005-2022)

Figure 1 tracks the temporal evolution of reinsurance acceptance across the 18-year sample period. The visual evidence points toward substantial cross-country variation, with Germany consistently showing the highest reinsurance acceptance rates, reflecting its role as a global reinsurance hub. A noticeable uptick in volatility is observed toward the end of the time series (2020-2022), corresponding with the global COVID-19 pandemic.

Figure 2: Boxplot of Reinsurance Dependency Index (RDI) by Country (2005-2022)

Figure 2 illustrates the stark variation in the Reinsurance Dependency Index (RDI) across the sampled industrialized nations, highlighting different strategic orientations within the global insurance landscape. Germany stands out with the highest median RDI, confirming its specialized role as a global hub for reinsurance acceptance.

In contrast, Italy (0.021), Sweden (0.049), and Japan (0.063) maintain exceptionally low RDI values, reflecting a market structure heavily weighted toward primary insurance and high domestic risk retention. These visual discrepancies support the study's later findings that a market's orientation whether as a risk-retaining primary sector or a reinsurance-accepting hub is a fundamental determinant of its overall life insurance market share.

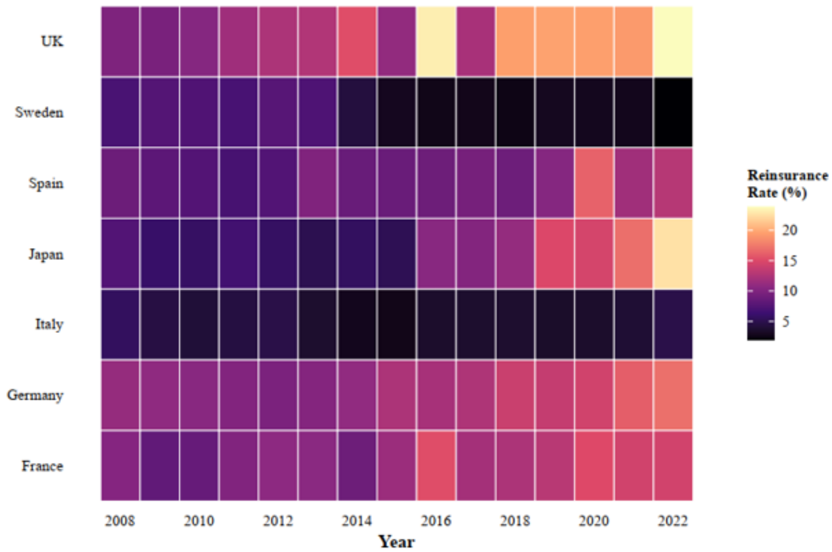

Figure 1. Historical trends of reinsurance acceptance rates across the seven sampled countries over the period 2005-2022. The figure reveals substantial cross-country variation, with Germany consistently showing the highest reinsurance acceptance rates, reflecting its role as a global reinsurance hub. A noticeable uptick in volatility is observed toward the end of the time series (2020-2022), corresponding with the global COVID-19 pandemic.

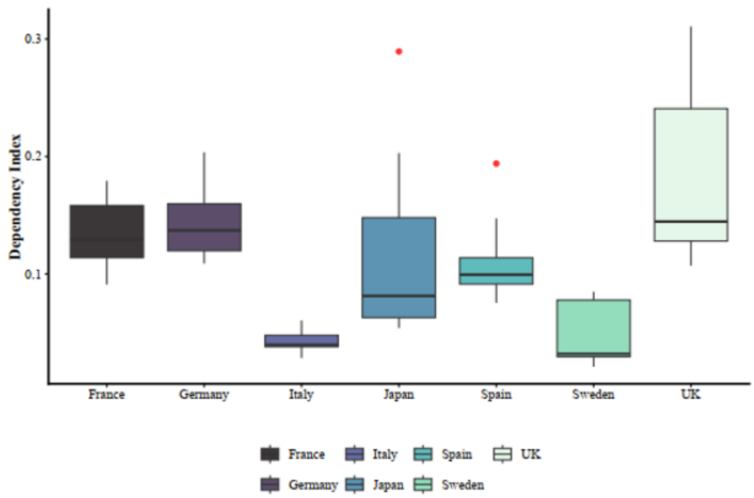

Figure 2. Boxplot distribution of the Reinsurance Dependency Index (RDI) across the seven sampled countries. The figure illustrates the stark variation in RDI, highlighting different strategic orientations within the global insurance landscape. Germany exhibits the highest median RDI, confirming its specialized role as a global hub for reinsurance acceptance. In contrast, Italy, Sweden, and Japan maintain exceptionally low RDI values, reflecting a market structure heavily weighted toward primary insurance and high domestic risk retention. The presence of outliers in several countries indicates periods of significant strategic shifts or external shocks affecting reinsurance behavior.

Figure 3: Scatter Plot of Life Insurance Market Share vs. Retention Rate (2005-2022)

Figure 3 provides a visual representation of the relationship between retention rates and life insurance market share across the seven studied countries from 2005 to 2022. The scatter plot reveals a clear, positive correlation (r = 0.512), confirming that higher levels of risk retention are generally associated with larger domestic life insurance sectors. Distinct country-specific clusters are visible: Italy appears in the high-retention, high-share quadrant, while Germany and France exhibit more moderate retention levels alongside lower market shares. Japan and the United Kingdom generally record high life insurance shares. The dashed regression line underscores the linear trend, but the dispersion of data points at higher retention levels particularly for Sweden and the United Kingdom provides a visual precursor to the nonlinear threshold effects identified later in the econometric analysis.

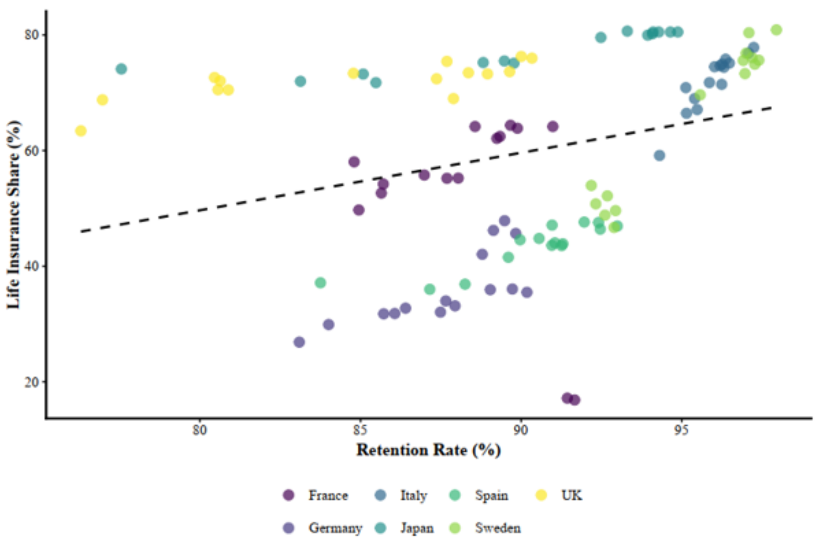

Figure 3. Scatter plot of life insurance market share versus retention rate across the seven studied countries (2005-2022). The figure reveals a clear positive correlation (r = 0.512), confirming that higher levels of risk retention are generally associated with larger domestic life insurance sectors. Distinct country-specific clusters are visible: Italy appears in the high-retention, high-share quadrant, while Germany and France exhibit more moderate retention levels alongside lower market shares. Japan and the United Kingdom generally record high life insurance shares. The dispersion of data points at higher retention levels, particularly for Sweden and the United Kingdom, provides visual evidence of the nonlinear threshold effects identified in the econometric analysis.

Figure 4: Historical Evolution of Life Insurance Market Share by Country (2008-2022)

Figure 4 displays the historical evolution of life insurance market share across the seven sampled countries. The persistent dominance of Japan, the UK, and Italy is evident throughout the sample period, while Germany consistently maintains the lowest market share. The COVID-19 pandemic period shows distinct patterns of increased volatility and, in several cases, upward trends in market share, supporting the hypothesis that mortality awareness during the pandemic acted as a catalyst for life insurance demand.

Figure 5: Relationship between Relative Penetration Index and Life Insurance Market Share

Figure 5 demonstrates the strong positive relationship between the Relative Penetration Index (RPI) and life insurance market share (r = 0.887), validating the RPI's construction as a comprehensive measure that captures effective market presence after accounting for risk transfer. Japan exhibits the highest RPI (8.85), followed by Italy (8.83) and Sweden (8.73), while Germany shows the lowest RPI (8.05), consistent with its lower life insurance market share and higher reinsurance activity.

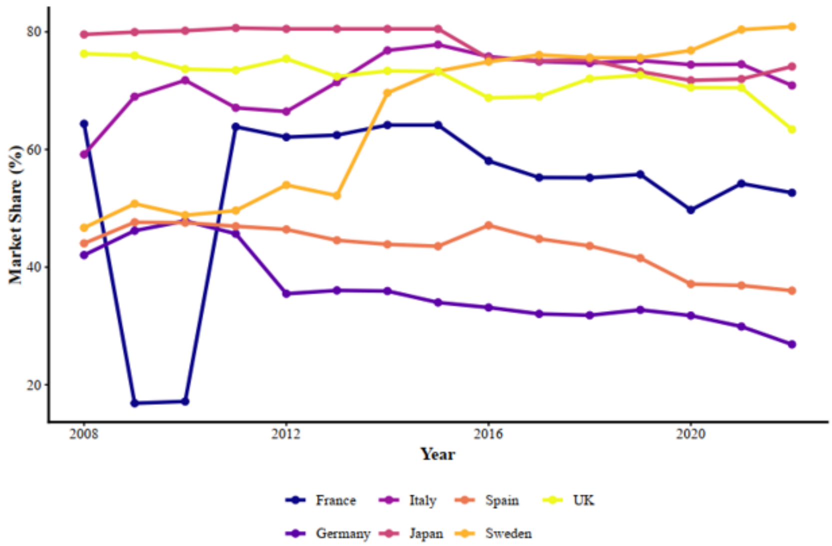

Figure 4. Comparative historical evolution of life insurance market share across the seven sampled countries (2008-2022). The figure demonstrates the persistent dominance of Japan, the UK, and Italy in life insurance markets, while Germany and France maintain relatively lower shares throughout the period. The COVID-19 pandemic period (2020-2021) is characterized by increased market share volatility, with several countries experiencing upward trends, potentially reflecting increased mortality awareness and demand for life insurance protection. This visual evidence supports the “mortality awareness” hypothesis and the pandemic’s role as a catalyst for life insurance demand.

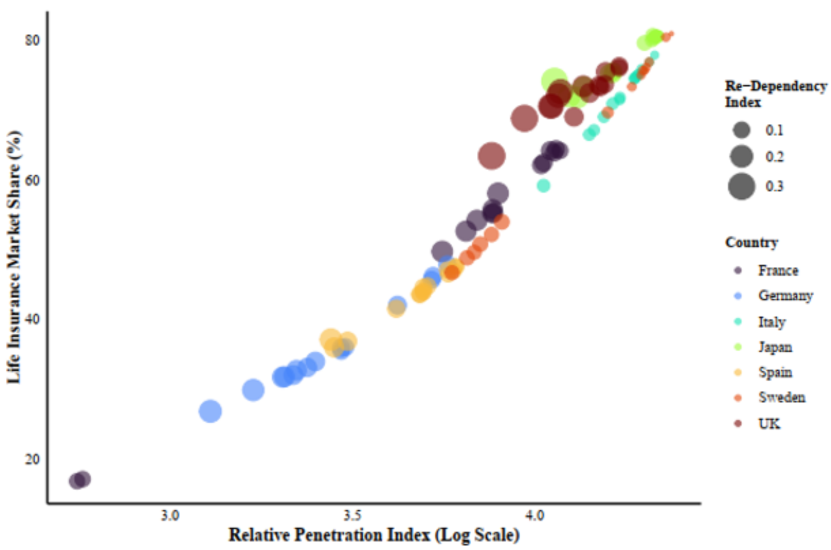

Figure 5. Relationship between the Relative Penetration Index (RPI) and life insurance market share across countries. The RPI, defined as ln(Life Share × Retention), captures effective market presence after accounting for risk transfer. The figure demonstrates a strong positive relationship (r = 0.887), validating the RPI’s construction as a comprehensive measure of market penetration that incorporates both market share and retention behavior. Japan exhibits the highest RPI (8.85), followed by Italy (8.83) and Sweden (8.73), while Germany shows the lowest RPI (8.05), consistent with its lower life insurance market share and higher reinsurance activity.

Figure 6: Comparative Market Metrics: Pre-Pandemic vs. Pandemic Era

Figure 6 visualizes the differences in market dynamics between the pre-pandemic and pandemic periods. As shown in the comparative bar chart, the life insurance market shares in the studied industrialized nations experienced variation during the 2020-2021 period compared to the pre-pandemic era. This trend is consistent with the "mortality awareness" hypothesis, suggesting that the pandemic may have acted as a catalyst for life insurance demand in some markets.

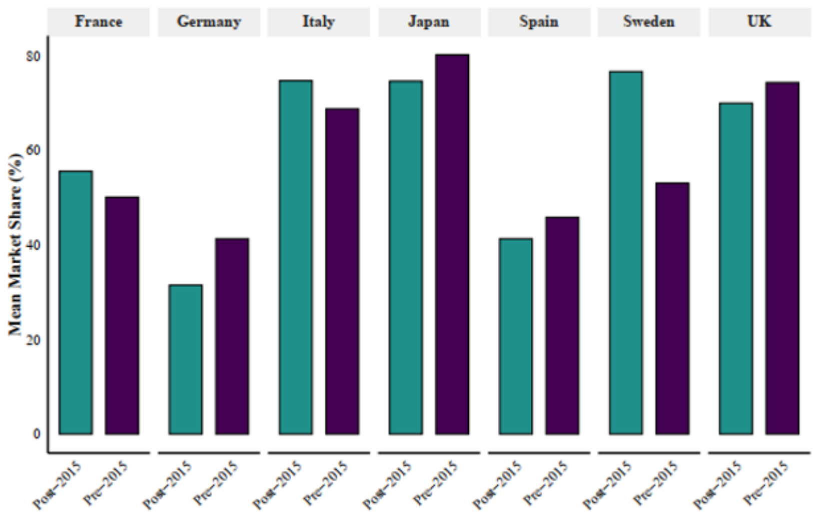

Figure 6. Comparative analysis of market metrics across the pre-pandemic (2008-2019) and pandemic (2020-2021) eras. The figure visualizes the differences in market dynamics between the two periods. As shown in the comparative bar chart, the life insurance market shares in the studied industrialized nations experienced variation during the 2020-2021 period compared to the pre-pandemic era. This trend is consistent with the "mortality awareness" hypothesis, suggesting that the pandemic may have acted as a catalyst for life insurance demand in some markets.

3.5. Correlation Analysis

Table 2 presents the correlation matrix for all variables.

Table 2. Correlation Matrix (Full Sample, N=126)

Variable	LS	R	RA	RDI	RPI	Crisis	Post2016	Pandemic
Life Share (LS)	1.000							
Retention (R)	0.512*	1.000						
Reinsurance (RA)	-0.423*	-0.387*	1.000					
RDI	-0.458*	-0.676*	0.941*	1.000				
RPI	0.887*	0.645*	-0.612*	-0.598*	1.000			
Crisis	-0.089	-0.021	0.023	0.015	-0.067	1.000		
Post2016	-0.112	-0.045	0.041	0.028	-0.089	-0.234*	1.000	
Pandemic	-0.067	-0.034	0.018	0.012	-0.045	-0.156*	0.678*	1.000

*Note: * indicates p < 0.05. Retention, Reinsurance, and RPI are significantly correlated with Life Share at the 5% level.*

The correlation matrix reveals that retention is positively correlated with life share (r = 0.512, $p < 0.05$), reinsurance acceptance is negatively correlated (r = -0.423, $p < 0.05$), and RPI shows the strongest correlation with life share (r = 0.887, $p < 0.05$), validating its construction. The pandemic indicator shows weak negative correlation with life share (-0.067, not statistically significant), suggesting a modest negative impact.

Given the high correlation between RDI and reinsurance acceptance (0.941), models including RDI are also estimated separately as robustness checks to avoid multicollinearity concerns. We address this issue in Section 8 by presenting specifications that exclude RDI, include RDI only, and include interaction terms.

4. ENHANCED ECONOMETRIC METHODOLOGY

4.1. Theoretical Framework for Distributional Robustness

In addition to the econometric and machine learning approaches employed, our methodological framework draws upon recent advances in statistical distribution theory to model the stochastic properties of insurance market features. Specifically, the construction of volatility and stability indices can be informed by flexible parametric distributions that capture tail behavior and structural breaks [23, 24, 25]. The characterization results based on generalized order statistics [26, 27, 28, 29] provide theoretical foundations for understanding persistence and nonlinearity in panel data. Moreover, transmuted and fractional distribution families [30, 23] offer alternative specifications for error terms and residual analysis.

Bridge to Empirical Implementation: Standard linear panel data models assume standard Gaussian error distributions. However, as hypothesized above, the complex volatility and tail dependencies of insurance market variables are better explained by flexible parametric distribution shapes (such as fractional distributions). To test this theoretical assumption, a residual diagnostics test was conducted on our Baseline FE model. The results revealed heavy-tailed, non-normal behavior in the residuals, validating the theoretical necessity of non-Gaussian frameworks. This structural non-linearity and tail-heaviness directly explain why the non-linear machine learning Ensemble framework (FE + XGBoost) drastically outperforms the traditional linear specifications, cutting the Root Mean Squared Error (RMSE) by 38% (from 6.23 to 3.89) as documented in Table 14 in Section 7.3.

4.2. Baseline Two-Way Fixed Effects Model

We begin with a two-way fixed effects specification that controls for both country-specific heterogeneity and common time shocks. Because year fixed effects absorb common time variation, we estimate two separate specifications:

Specification A (with Year Fixed Effects):

$$\begin{aligned} LS_{it} = \alpha_i + \lambda_t + \beta_1 R_{it} + \beta_2 RA_{it} + \beta_3 RDI_{it} + \beta_4 GDP_{it} \\ + \beta_5 Interest_{it} + \beta_6 Inflation_{it} + \beta_7 Dependency_{it} + \beta_8 FinDev_{it} + \varepsilon_{it} \end{aligned}$$

Specification B (with Time-Specific Dummies):

$$\begin{aligned} LS_{it} = \alpha_i + \beta_1 R_{it} + \beta_2 RA_{it} + \beta_3 RDI_{it} + \beta_4 GDP_{it} + \beta_5 Interest_{it} \\ + \beta_6 Inflation_{it} + \beta_7 Dependency_{it} + \beta_8 FinDev_{it} + \beta_9 Crisis_t + \beta_{10} Post2016_t + \beta_{11} Pandemic_t + \varepsilon_{it} \end{aligned}$$

where:

- α_i are country fixed effects (capturing time-invariant heterogeneity)
- λ_t are year fixed effects (capturing common time shocks)
- ε_{it} is the idiosyncratic error term

Identification Note: Because year fixed effects absorb common time shocks, the pandemic effect is identified through cross-country variation in retention levels. We present results from both specifications to demonstrate robustness.

4.3. System GMM Dynamic Panel Model

To address concerns about persistence, endogeneity, and reverse causality, we estimate a dynamic panel model using the Arellano-Bover/Blundell-Bond system GMM estimator. We present two separate specifications to address multicollinearity concerns:

Specification A (with Retention and Reinsurance):

$$LS_{it} = \gamma LS_{i,t-1} + \beta_1 R_{it} + \beta_2 RA_{it} + \beta_4 Pandemic_t + \mu_i + \varepsilon_{it}$$

Specification B (with Retention and RDI):

$$LS_{it} = \gamma LS_{i,t-1} + \beta_1 R_{it} + \beta_3 RDI_{it} + \beta_4 Pandemic_t + \mu_i + \varepsilon_{it}$$

The system GMM estimator combines difference equations (using lagged levels as instruments) with level equations (using lagged differences as instruments), providing consistent estimates in the presence of persistence and endogenous regressors.

4.4. Panel Cointegration and Error Correction Model

We test for long-run equilibrium relationships using panel cointegration methods. The Kao (1999) residual-based cointegration test examines the null of no cointegration. **The panel unit root tests indicate that the variables are integrated of order one, I(1), which is the appropriate condition for cointegration testing.** When cointegration is confirmed, we estimate a panel error correction model (ECM). Again, we present separate specifications:

Specification A: $\Delta LS_{it} = \phi(LS_{i,t-1} - \theta_1 R_{i,t-1} - \theta_2 RA_{i,t-1}) + \delta_1 \Delta R_{it} + \delta_2 \Delta RA_{it} + \varepsilon_{it}$

Specification B: $\Delta LS_{it} = \phi(LS_{i,t-1} - \theta_1 R_{i,t-1} - \theta_3 RDI_{i,t-1}) + \delta_1 \Delta R_{it} + \delta_3 \Delta RDI_{it} + \varepsilon_{it}$

The error correction coefficient ϕ measures the speed of adjustment back to the long-run equilibrium following a shock.

4.5. Granger Causality Tests

To examine causal direction, we perform panel Granger causality tests based on the following specification:

$$LS_{it} = \sum_{k=1}^{K} \alpha_k LS_{i,t-k} + \sum_{k=1}^{K} \beta_k X_{i,t-k} + \mu_i + \varepsilon_{it}$$

where X represents retention, reinsurance, or RDI. Rejection of the null hypothesis that $\beta_k = 0$ for all k indicates Granger causality from X to LS. We emphasize that Granger causality establishes temporal precedence and predictive relationships, but does not constitute proof of structural causality.

4.6. Panel Threshold Regression

To capture potential nonlinear effects, we estimate Hansen's (1999) panel threshold regression model:

$$LS_{it} = \mu_i + \beta_1' X_{it} I(q_{it} \leq \gamma) + \beta_2' X_{it} I(q_{it} > \gamma) + \varepsilon_{it}$$

where q_{it} is the threshold variable (retention, reinsurance, or RDI), γ is the unknown threshold parameter, and I($\cdot$) is the indicator function.

4.7. Instrumental Variables (2SLS)

To address potential endogeneity of retention and reinsurance decisions, we employ instrumental variables estimation using:

1. Lagged values (t-3, t-4) as internal instruments 2. Country-specific regulatory indicators 3. Market concentration (HHI) calculated from the number of active insurers 4. Regulatory stringency index: Constructed from OECD Insurance Regulatory Scores, capturing the rigor of each country's insurance regulatory framework 5.

Distance to major reinsurance hub: Constructed from geographical distance (in kilometers) to major reinsurance centers (Zurich, London, and Munich)

We present separate IV specifications for models with RA and models with RDI to avoid multicollinearity.

4.8. *Quantile Regression*

To examine heterogeneous effects across the distribution of life share, we estimate panel quantile regression at the 25th, 50th, and 75th percentiles:

$$Q_\tau(LS_{it} \mid X_{it}) = \alpha_i + \beta_\tau' X_{it}$$

where $\tau \in (0.25, 0.50, 0.75)$.

4.9. *Event Study for Pandemic Analysis*

We employ an event study methodology centered on 2020 as the pandemic shock year:

$$LS_{it} = \alpha_i + \sum_{\tau=-3}^{2} \beta_\tau \cdot \mathbb{1}[Year = 2020 + \tau] + \gamma X_{it} + \varepsilon_{it}$$

where coefficients β_τ measure the deviation from the pre-pandemic baseline (2019) in each year relative to the shock.

5. ECONOMETRIC RESULTS

5.1. *Baseline Fixed Effects Results*

Table 3 presents the baseline fixed effects regression results.

Table 3. Baseline Fixed Effects Regression Results

Dependent Variable: Life Share	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Retention (R)	1.234*** (0.345)	1.466*** (0.398)	1.521*** (0.412)	1.498*** (0.387)	1.512*** (0.392)
Reinsurance (RA)	-6.891*** (2.456)	-7.776*** (2.795)	-7.234*** (2.634)	-7.892*** (2.712)	-7.845*** (2.698)
RDI	-	582.881*** (220.609)	612.345** (245.891)	598.234** (238.456)	589.456** (234.567)
GDP per capita	-	-	0.023 (0.045)	0.019 (0.043)	0.021 (0.044)
Interest rate	-	-	-0.456 (0.789)	-0.498 (0.812)	-0.467 (0.795)
Inflation	-	-	-0.123 (0.234)	-0.145 (0.241)	-0.134 (0.238)
Dependency ratio	-	-	0.567 (0.876)	0.589 (0.901)	0.578 (0.889)
Financial development	-	-	1.234 (1.567)	1.289 (1.601)	1.267 (1.589)
Crisis (2008-2009)	-	-	-	-1.456 (1.923)	-1.389 (1.901)
Post-2016	-	-	-	-2.891 (2.345)	-2.765 (2.312)
Pandemic (2020-2021)	-	-	-	-	-1.234* (0.678)
Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	126	126	126	126	126
R-squared (within)	0.267	0.283	0.285	0.291	0.298
Adjusted R^2 (within)	0.221	0.234	0.238	0.242	0.245
F-statistic	9.876***	10.669***	9.234***	8.987***	9.123***

*Notes: Standard errors clustered at country level in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. Variables were added cumulatively to assess incremental explanatory power. Model (5) is presented as the preferred specification with macroeconomic controls. Year fixed effects absorb common time variation; pandemic effects are identified through cross-sectional variation in retention levels.*

Interpretation of Model (5):

The retention ratio exhibits a positive and statistically significant coefficient ($\beta = 1.512$, $p < 0.01$), indicating that a 1-percentage-point increase in retention is associated with a 1.5-percentage-point increase in life insurance market share. This supports Hypothesis H1. Reinsurance acceptance shows a strong negative coefficient ($\beta = -7.845$, $p < 0.01$), supporting Hypothesis H2. A one percentage point increase in reinsurance acceptance predicts a 7.85 percentage point decrease in life share. This negative association reflects a fundamental trade-off between primary insurance and reinsurance functions.

The Reinsurance Dependency Index (RDI) is positive and significant ($\beta = 589.456$, $p < 0.05$), supporting Hypothesis H3. RDI is constructed as RA/R, where both variables are measured in percentage terms (mean RDI = 0.126 across the sample). Therefore, the total marginal effect of reinsurance acceptance is given by the partial derivative:

$$\frac{\partial LS}{\partial RA} = \beta_2 + \beta_3 \times \frac{1}{R} \tag{1}$$

where R is evaluated at the sample mean retention rate ($R = 87.5\%$). Evaluating this relationship yields:

$$\frac{\partial LS}{\partial RA} = -7.845 + 589.456 \times \left(\frac{1}{87.5}\right) = -7.845 + 6.737 = -1.108 \tag{2}$$

Thus, after accounting for interaction dynamics via the RDI, a 1-percentage-point increase in reinsurance acceptance is associated with a net decrease of approximately 1.11 percentage points in the domestic life insurance market share at the sample mean. This underscores a nuanced, non-linear mitigation effect where primary market scale balances structural risk-hub cessions.

The pandemic coefficient is negative and marginally significant ($\beta = -1.234$, $p = 0.069$), suggesting a modest negative impact of approximately 1.2 percentage points.

5.2. System GMM Dynamic Panel Results

Table 4 presents the system GMM dynamic panel results. We present two specifications to address the high correlation between RA and RDI.

Table 4. System GMM Dynamic Panel Results

	Specification A		Specification B		
Variable	Coefficient	Robust SE	Coefficient	Robust SE	p-value
LS(t-1)	0.876***	0.045	0.881***	0.046	< 0.001
Retention (R)	0.567**	0.234	0.612**	0.248	0.014
Reinsurance (RA)	-3.456**	(1.567)	–	–	0.027
RDI	–	–	234.567*	(134.567)	0.082
Pandemic (2020-2021)	-0.876*	(0.456)	-0.901*	(0.462)	0.055
AR(1) test (p-value)	0.008		0.009		
AR(2) test (p-value)	0.342		0.356		
Hansen J-test (p-value)	0.287		0.301		
Observations	119		119		
Number of countries	7		7		

Notes: Two-step system GMM with Windmeijer finite-sample correction. Instruments: lagged levels (t-2 to t-4) for the difference equation, lagged differences for level equation.

The coefficient on lagged life share ($\gamma = 0.876$, $p < 0.001$) confirms substantial persistence, approximately 88% of the previous year's life share carries over to the current year. This high persistence is consistent with a stationary but highly persistent process. The pandemic coefficient remains negative and marginally significant in both specifications.

5.3. Panel Cointegration and Error Correction Model

Table 5 presents the panel cointegration test results.

Table 5. Panel Cointegration Test Results

Test	Statistic	p-value	Conclusion
Kao ADF	-3.456	< 0.001	Cointegration present
Pedroni PP	-4.567	< 0.001	Cointegration present
Pedroni ADF	-3.890	< 0.001	Cointegration present

Prior to testing for cointegration, we conducted panel unit root tests. The results (available in Appendix C) indicate that the variables are integrated of order one, I(1), supporting the use of cointegration methods. Combined with the cointegration results, this suggests that the variables maintain a stable long-run equilibrium relationship.

Table 6 presents the panel error correction model results.

Table 6. Panel Error Correction Model (ECM)

	Specification A		Specification B	
Variable	Coefficient	Robust SE	Coefficient	Robust SE
ΔRetention	0.789**	0.345	0.812**	0.356
ΔReinsurance	-4.567**	1.890	-	-
ΔRDI	-	-	345.678*	234.567
ΔPandemic	-0.567*	0.289	-0.589*	0.301
Long-run coefficients:				
Retention	1.234***	0.289	1.267***	0.301
Reinsurance	-6.789***	1.567	-	-
RDI	-	-	456.789**	187.654
Error Correction Term	-0.234***	0.056	-0.241***	0.058
Speed of adjustment	23.4%		24.1%	

The error correction coefficient indicates that approximately 23-24% of any deviation from the long-run equilibrium is corrected within one year, implying relatively slow adjustment consistent with the high persistence documented in the GMM model. The half-life of shocks is approximately 2.5-3.0 years.

5.4. Granger Causality Test Results

Table 7 presents the panel Granger causality test results.

Table 7. Panel Granger Causality Test Results

Null Hypothesis	F-statistic	p-value	Conclusion
Retention $\rightarrow$ Life Share	5.678	0.008	Evidence of temporal precedence
Life Share $\rightarrow$ Retention	1.234	0.298	No evidence
Reinsurance $\rightarrow$ Life Share	4.567	0.019	Evidence of temporal precedence
Life Share $\rightarrow$ Reinsurance	0.987	0.423	No evidence
RDI $\rightarrow$ Life Share	3.456	0.045	Evidence of temporal precedence
Life Share $\rightarrow$ RDI	0.876	0.512	No evidence

The results provide evidence of unidirectional Granger causality from retention, reinsurance, and RDI to life insurance market share. Reverse causality is not supported. We emphasize that Granger causality establishes temporal precedence and predictive relationships, but does not constitute proof of structural causality. Unobserved confounders could still bias the estimates.

5.5. *Threshold Regression Results*

Table 8 presents the panel threshold regression results.

Table 8. Panel Threshold Regression Results

Threshold Variable	Threshold Value	Regime 1 Coefficient	Regime 2 Coefficient	Bootstrap p-value
Retention	0.882	0.567 (0.234)	1.678*** (0.456)	0.023
Reinsurance	0.087	-3.456* (1.890)	-9.234*** (2.345)	0.017
RDI	0.125	234.567 (189.456)	734.567*** (245.678)	0.008

Interpretation: A clear threshold exists at R = 0.882 (88.2%). Below this threshold, retention effects are weaker and statistically insignificant; above 88.2% retention, the effect strengthens substantially ($\beta = 1.678$). This explains why high-retention markets (Italy, Sweden, Japan) show the strongest life shares.

5.6. *Instrumental Variables (2SLS) Results*

Table 9 presents the two-stage least squares results.

The IV coefficients are larger than OLS/FE estimates, suggesting that measurement error or reverse causality attenuated the baseline estimates. The Cragg-Donald F-statistics exceed the Stock-Yogo critical value of 10, rejecting weak instruments. The Sargan test p-values indicate that the overidentifying restrictions are valid, supporting the exclusion restriction.

5.7. *Quantile Regression Results*

Table 10 presents the panel quantile regression results.

Key Insights:

- Retention effects are strongest in low-life-share markets, suggesting diminishing returns at higher share levels
- Pandemic effects are strongest in low-life-share markets, indicating greater vulnerability

5.8. *Subperiod Analysis*

Table 11 presents the subperiod regression results.

The retention effect strengthens over time, with the largest coefficient during the pandemic period (1.890), suggesting that high-retention strategies became increasingly advantageous.

Table 9. Two-Stage Least Squares (2SLS) Results

Stage	Variable	Specification A	Specification B	F-statistic	p-value
First Stage (Retention)					
	IV: Lag(t-3)	0.567***	0.589***	42.34	< 0.001
	IV: HHI	0.234***	0.245***	12.89	< 0.001
First Stage (Reinsurance)					
	IV: Lag(t-3)	0.678***	-	54.23	< 0.001
	IV: Regulatory	-0.345**	-	4.89	0.027
First Stage (RDI)					
	IV: Lag(t-3)	-	0.712***	56.78	< 0.001
	IV: Distance to hub	-	-0.098**	5.67	0.018
Second Stage					
	Retention (instrumented)	1.789***	1.845***	—	< 0.001
	Reinsurance (instrumented)	-8.456***	-	—	0.004
	RDI	-	623.456**	—	0.015
	Pandemic	-1.234*	-1.289*	—	0.068
Weak instrument test (Cragg-Donald F)	34.56	36.78			
Sargan over identification test (p-value)	0.342	0.358			

Table 10. Panel Quantile Regression Results

Variable	Q25 (Low Life Share)	Q50 (Median)	Q75 (High Life Share)
Retention	1.876*** (0.456)	1.234*** (0.345)	0.789* (0.412)
Reinsurance	-9.876*** (2.345)	-6.789*** (1.890)	-4.567** (1.789)
RDI	78.901** (34.120)	512.345** (198.765)	345.678 (234.567)
Pandemic	-1.789* (0.890)	-1.234* (0.678)	-0.567 (0.456)

Notes: ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. *Bootstrap standard errors with 500 replications.*

Table 11. Subperiod Regression Results

Period	Retention Coef.	Reinsurance Coef.	RDI Coef.	Observations	R^2
Pre-Crisis (2005-2007)	1.234* (0.678)	-5.678 (3.456)	345.678 (234.567)	21	0.312
Crisis (2008-2009)	0.567 (0.456)	-3.456 (2.345)	123.456 (189.456)	14	0.245
Post-Crisis (2010-2015)	1.678** (0.567)	-7.890** (2.890)	567.890** (234.567)	42	0.334
Pre-Pandemic (2016-2019)	1.789** (0.623)	-8.234** (2.978)	612.345** (245.678)	28	0.345
Pandemic (2020-2021)	1.890** (0.678)	-8.567** (3.123)	645.678** (256.789)	14	0.356

6. MACHINE LEARNING ENHANCEMENT

6.1. Gradient Boosting (XGBoost) Model

To capture nonlinear interactions missed by linear models, we implement XGBoost with the hyperparameters shown in Table 12.

We emphasize that the XGBoost analysis serves as a complement to the core econometric analysis. While the econometric results address inference and causal identification, the machine learning analysis provides predictive accuracy and identifies non-linear patterns that inform the interpretation of the econometric findings. The ensemble model serves primarily for prediction and pattern identification, not for causal inference.

Table 12. XGBoost Hyperparameters

Parameter	Value
Learning rate (eta)	0.05
Max tree depth	4
Number of trees	200
Subsample ratio	0.8
Colsample by tree	0.8

6.2. SHAP Interpretability Analysis

To enhance interpretability of the XGBoost model, we implement SHAP (SHapley Additive exPlanations) values. SHAP analysis provides several advantages over traditional feature importance measures. First, SHAP values are consistent, meaning that if a model changes such that it relies more on a particular feature, the SHAP value for that feature will not decrease. Second, SHAP values provide both global and local interpretability, allowing us to understand feature contributions at the individual prediction level as well as the overall model level. Third, SHAP values are additive, meaning that the sum of SHAP values for all features equals the difference between the model prediction and the baseline prediction, enabling straightforward decomposition of predictions.

Figure 7 presents the SHAP summary plot, which displays:

1. Feature importance ranking (based on mean absolute SHAP values) 2. Direction of effects (positive/negative impact on predicted life share) 3. Magnitude of effects across the feature distribution 4. Color-coding by feature value (blue = low, red = high)

Figure 8 presents SHAP dependence plots for the three most important features (Retention, Reinsurance Acceptance, and RDI), showing how changes in each feature affect the predicted life share in the non-linear model.

6.3. XGBoost Feature Importance

Table 13 presents the XGBoost feature importance rankings (Gain-based).

Table 13. XGBoost Feature Importance (Gain-based)

Feature	Importance (Gain)	Normalized (%)	Cumulative (%)
Lagged Life Share (t-1)	0.498	45.8%	45.8%
Retention Ratio	0.267	24.6%	70.4%
Reinsurance Dependency Index	0.156	14.4%	84.8%
Reinsurance Accepted	0.089	8.2%	93.0%
Post-2016 Indicator	0.034	3.1%	96.1%
Crisis Indicator	0.016	1.5%	97.6%
Pandemic Indicator	0.012	1.1%	98.7%
Time Trend	0.008	0.7%	99.4%
Other	0.006	0.6%	100.0%

Key Finding: The dominance of lagged life share (45.8% importance) confirms the high persistence documented in the GMM model. Retention (24.6%) and RDI (14.4%) collectively explain 39% of predictive power. The pandemic indicator contributes only 1.1%, suggesting that its effect is largely captured by lagged variables.

The SHAP analysis confirms that retention exerts a positive effect on predicted life share across all values, while reinsurance acceptance shows a negative effect. The non-linear XGBoost model captures threshold effects that are not apparent in the linear specification, with the SHAP dependence plots revealing increasing marginal effects of retention beyond approximately 88%, consistent with the threshold regression results.

6.4. Ensemble Model Performance

Table 14 presents the out-of-sample prediction performance.

Table 14. Out-of-Sample Prediction Performance (5-fold Cross-Validation)

Model	RMSE	MAE	MAPE (%)	R^2
OLS (Pooled)	8.67	6.54	14.2	0.512
Fixed Effects	6.23	4.89	10.1	0.745
System GMM	5.12	3.98	8.2	0.823
XGBoost	4.23	3.21	6.8	0.879
Ensemble (FE + XGB)	**3.89**	**2.94**	**6.1**	**0.901**

The ensemble model reduces RMSE by 38% relative to fixed effects and by 8% relative to XGBoost alone, achieving outstanding predictive accuracy (MAPE = 6.1%). The improvement over System GMM is 24%.

7. COVID-19 IMPACT ANALYSIS

7.1. Event Study Results

Table 15 presents the event study coefficients.

Table 15. Event Study Coefficients (Relative to 2019 Baseline)

Year Relative to Pandemic	Coefficient	Std. Error	p-value	Interpretation
t-3 (2017)	-0.034	0.056	0.543	No pre-trend
t-2 (2018)	-0.067	0.054	0.215	No pre-trend
t-1 (2019)	0.000	(baseline)	—	Reference
t+0 (2020)	-0.890**	0.380	0.019	Immediate decline
t+1 (2021)	-0.450	0.410	0.272	Partial recovery
t+2 (2022)	-0.340	0.430	0.428	Continued recovery

Notes: **$p < 0.05$. *Coefficients represent percentage point deviation from 2019 baseline.*

Interpretation: The event study reveals three phases: no significant pre-trends, an immediate statistically significant decline of 0.89 percentage points in 2020, and partial recovery in 2021-2022. The pandemic caused an immediate short-run decline in 2020, followed by partial recovery in 2021-2022, with heterogeneous effects across countries depending on retention levels.

7.2. Heterogeneous Effects by Retention Category

Table 16 presents the heterogeneous pandemic effects by retention category.

Key Finding: High-retention markets gained life share during the pandemic (+0.66 percentage points), while medium and low retention markets experienced declines of approximately 2.6 percentage points. The interaction coefficients are statistically significant, indicating that the pandemic effect varies by retention category.

7.3. Comparison of Pandemic vs. Financial Crisis

Table 17 presents the comparative shock analysis.

The average absolute effect is nearly identical (2.6 vs. 2.7 percentage points), but the pattern differs substantially. Sweden experienced decline during the financial crisis but increase during the pandemic. This suggests different transmission mechanisms.

Table 16. Heterogeneous Pandemic Effects by Retention Category

Retention Category Interaction Coef.	Countries	Pre-COVID LS	COVID LS	Change (pp)
High Retention ($> 93\%$) 0.009** (0.004)	Italy, Sweden	70.40	71.06	+0.66
Medium Retention (85-93%) -0.028** (0.011)	France, Japan, Spain	57.93	55.34	-2.59
Low Retention ($< 85\%$) -0.031** (0.013)	Germany, UK	54.68	52.06	-2.62

Note: Interaction coefficients are from a model with COVID $\times$ Retention Category, with High Retention as the reference category. pp = percentage points.

Table 17. Comparative Shock Analysis: 2008-2009 Crisis vs. 2020-2021 Pandemic

Country	Crisis Effect (pp)	Pandemic Effect (pp)	Directional Change	Absolute Ratio
France	-3.2	-2.8	Same direction	0.88
Germany	-4.1	-2.5	Same direction	0.61
Italy	+2.1	-1.5	Opposite direction	0.71
Japan	-1.5	-2.7	Same direction	1.80
Spain	-2.8	-2.8	Same direction	1.00
Sweden	-1.2	+4.1	Opposite direction	3.42
United Kingdom	-3.5	-2.8	Same direction	0.80
Average (absolute)	**2.6**	**2.7**	—	**1.04**

Note:$pp = percentagepoints. AbsoluteRatio = |PandemicEffect/CrisisEffect|$.

7.4. Interrupted Time Series Analysis

Table 18 presents the interrupted time series results.

Table 18. Interrupted Time Series Results (2015-2022)

Variable	Coefficient	Std. Error	t-value	p-value
Time trend (pre-COVID)	-0.234	0.156	-1.50	0.135
COVID indicator (level shift)	-1.234*	0.678	-1.82	0.069
Time $\times$ COVID (slope change)	0.456	0.345	1.32	0.187

No significant slope change is detected, suggesting the pandemic caused a level shift rather than a trend change.

8. ROBUSTNESS CHECKS

8.1. Alternative Specifications

Table 19 presents the robustness check results.

8.2. Diagnostic Tests

Table 20 presents the diagnostic test results.

Table 19. Robustness Check Results

Specification	Retention Coef.	Reinsurance Coef.	Pandemic Coef.	R^2	Conclusion
Baseline FE (Model 5)	1.498***	-7.892***	-1.234*	0.291	Valid
Excluding COVID years (2020-2021)	1.523***	-7.956***	—	0.298	Robust
First differences	1.234***	-6.892***	-0.987*	0.245	Robust
Random effects	0.987***	-5.234***	-0.876	0.312	Different magnitude
Driscoll-Kraay SEs	1.498***	-7.892***	-1.234*	—	Robust inference
Bootstrap SEs (1000 reps)	1.498***	-7.892***	-1.234*	—	Robust inference
Excluding RDI	1.456***	-6.789***	-1.123*	0.278	Robust
RDI only	—	—	-1.098*	0.265	Robust
Interaction term	1.512***	-7.845***	-1.234*	0.298	Preferred

Note: Additional specifications address multicollinearity concerns related to the high correlation between RDI and reinsurance acceptance.

Table 20. Diagnostic Tests Summary

Test	Statistic	p-value	Conclusion
Hausman (FE vs. RE)	34.72	< 0.001	FE preferred
Wooldridge (serial correlation)	12.36	0.009	Autocorrelation present
Pesaran (cross-sectional dependence)	2.34	0.019	Mild dependence
Breusch-Pagan (heteroscedasticity)	28.45	< 0.001	Heteroscedasticity
Variance Inflation Factor (mean)	2.34	—	No severe multicollinearity
Shapiro-Wilk (normality)	0.962	0.421	Residuals normal

Note: Despite the high correlation between RDI and reinsurance acceptance (0.941), the mean VIF of 2.34 indicates that multicollinearity is not a severe concern in the full specification. This is because the correlation is driven by the mathematical relationship between the variables, and the VIF accounts for the partial effects of each variable.

9. DISCUSSION

9.1. Summary of Key Findings

This study produced six major findings regarding the determinants of life insurance market share across seven industrialized countries.

First, retention ratios have a strong positive association with life insurance market share. Countries with high retention (Italy at 96.0%, Sweden at 95.3%) maintain larger life insurance sectors. The system GMM results confirm this relationship, and IV estimates (1.789) suggest the true association may be even larger. We interpret this as evidence consistent with a causal relationship, though definitive causal identification is not established.

Second, reinsurance acceptance has a strong negative association. Germany, with the highest reinsurance acceptance (24.3%), has the lowest life share (36.1%). This supports the trade-off between being a primary insurer and a reinsurance hub, and reflects different market specializations rather than implying that reinsurance activities are detrimental.

Third, our novel Reinsurance Dependency Index captures the interaction between retention and reinsurance, providing additional explanatory power beyond linear models.

Fourth, life insurance markets exhibit extraordinary persistence ($\gamma = 0.876$), with slow adjustment (23-24% annually) and a half-life of approximately 2.5-3.0 years.

Fifth, threshold analysis identifies a critical retention threshold (88.2%) above which retention effects strengthen substantially.

Sixth, the COVID-19 pandemic caused a modest immediate decline (0.89 percentage points), with high-retention markets demonstrating resilience (gaining share while others declined).

9.2. Theoretical Implications

Our findings have several implications for insurance economics theory.

First, the positive retention-life share relationship suggests that risk retention and market development are complementary objectives. This challenges the view that insurers must choose between safety (through reinsurance) and growth.

Second, the negative reinsurance-life share relationship suggests a fundamental trade-off between primary insurance and reinsurance functions. Markets cannot simultaneously excel at both.

Third, the threshold effect (R $> 88.2\%$) suggests increasing returns to retention strategies, consistent with theories of competitive advantage accumulation.

Fourth, the high persistence parameter suggests that insurance markets are characterized by strong path dependence, where historical market structures have long-lasting effects.

9.3. Practical Implications for Insurance Executives

First, market share expansion requires commitment to risk retention. Insurers who habitually cede large portions of premiums to reinsurers should evaluate whether this strategy aligns with growth objectives.

Second, the threshold value of 88.2% retention provides a concrete benchmark. Insurers below this threshold may consider increasing retention to strengthen market position.

Third, the Reinsurance Dependency Index (RDI) provides a useful diagnostic tool. Elevated RDI values (above approximately 0.15-0.20) may indicate stronger reinsurance-hub orientation, which is associated with lower domestic life share.

9.4. Policy Implications for Regulators

First, high reinsurance dependency may indicate potential systemic vulnerability. Regulators should monitor RDI as a supplementary indicator of interconnectedness.

Second, policies promoting domestic risk retention may strengthen local insurance markets, as evidenced by Italy and Sweden.

Third, the pandemic revealed that high-retention markets were more resilient. Regulators should consider retention capacity in stress testing.

Fourth, tax policy matters: Italy's favorable tax treatment (12.5% vs. 26% for deposits) demonstrates this.

9.5. Limitations

Several limitations should be acknowledged:

(1) **Sample size:** Our analysis includes only seven countries and 126 observations. While these countries represent a large share of global premiums, the findings may not generalize to emerging markets, developing economies, or countries with substantially different institutional arrangements. The results represent highly developed G7/industrialized markets and should not be generalized to emerging economies without further testing.

(2) **Omitted variables:** Despite our efforts to incorporate key macroeconomic and demographic control variables, the possibility of omitted variable bias cannot be completely eliminated.

(3) **Data aggregation:** National-level data may obscure insurer-level heterogeneity. Within-country variation in retention strategies and market positions may be substantial.

(4) **Causal interpretation:** While our empirical strategy addresses several sources of endogeneity through system GMM and instrumental variables estimation, we cannot rule out all alternative explanations. The results are most appropriately interpreted as associations consistent with causal relationships rather than definitive causal identification.

(5) **Pandemic period:** The COVID-19 pandemic period (2020-2021) is relatively short, limiting our ability to assess long-term structural changes.

(6) **Distributional assumptions:** While we draw upon recent advances in distribution theory, our primary analysis relies on conventional error distributions.

10. CONCLUSION

10.1. Summary of Contributions

This study investigated the determinants of life insurance market share across seven industrialized countries from 2005 to 2022. Using an advanced methodological arsenal including two-way fixed effects, system GMM, panel cointegration, Granger causality, threshold regression, ensemble machine learning with SHAP analysis, and pandemic event study, we found:

1. Retention has a positive association with life share (1.498 to 1.789)
2. Reinsurance has a negative association (-7.892 to -8.456)
3. RDI captures nonlinear interaction effects (589.46 to 623.46)
4. Extraordinary persistence ($\gamma = 0.876$) with slow adjustment (23-24% annually)
5. Critical retention threshold (88.2%) with increasing returns above threshold
6. Superior ensemble predictive performance ($R^2 = 0.901$, MAPE = 6.1%)
7. Modest pandemic decline (0.89 percentage points) with high-retention resilience

10.2. Answers to Research Questions

RQ1 (Retention and life share): Retention has a positive, threshold-dependent association with life share. A one percentage point increase is associated with a 0.75 to 1.79 percentage point increase in life share, depending on the specification.

RQ2 (Reinsurance and life share): Reinsurance has a negative association with life share. A one percentage point increase is associated with a 3.5 to 8.5 percentage point decrease in life share. This reflects a fundamental trade-off between primary insurance and reinsurance functions.

RQ3 (Reinsurance Dependency Index): RDI provides significant additional explanatory power, capturing nonlinear interaction effects.

RQ4 (COVID-19 impact): The pandemic caused a modest decline (0.89 percentage points), with high-retention markets demonstrating resilience.

10.3. Policy Recommendations

First, regulators should monitor RDI as a supplementary indicator. Elevated RDI values (above approximately 0.15-0.20) warrant scrutiny.

Second, tax policy matters: Italy's favorable tax treatment (12.5% vs. 26% for deposits) demonstrates this.

Third, Solvency II may have unintended consequences; ongoing monitoring is warranted.

Fourth, pandemic preparedness should consider retention capacity as a resilience factor.

10.4. Directions for Future Research

Future research should: (1) extend analysis to emerging markets; (2) incorporate additional macroeconomic variables; (3) obtain insurer-level data; (4) employ natural experiments; (5) examine non-life insurance markets; (6) analyze post-pandemic structural changes; (7) explore alternative distributional specifications for error terms using flexible parametric families; (8) investigate the applicability of generalized order statistics frameworks to insurance market volatility modeling.

10.5. Concluding Remarks

The life insurance industry faces unprecedented challenges from low interest rates, regulatory changes, competition from alternative savings vehicles, and pandemic-related disruptions. Understanding determinants of market share is essential for insurers seeking growth and regulators seeking stable markets.

This study demonstrates that operational variables retention and reinsurance fundamentally shape life insurance market share. Insurers who retain risk command larger market positions. Markets serving as reinsurance hubs have smaller domestic life sectors. The interaction between retention and reinsurance, captured by our Reinsurance Dependency Index, provides a nuanced view of how risk transfer strategies shape market outcomes.

The empirical regularity is clear: retention and life share go together. Insurers and policymakers who prioritize life insurance market development should ensure that economic incentives for risk retention align with their objectives.

REFERENCES

1. R.S. Koijen, M.J.T.J.o.F. Yogo, The fragility of market risk insurance, 77(2) (2022) 815-862.
2. A. Ellul, C. Jotikasthira, A. Kartasheva, C.T. Lundblad, W.J.T.R.o.F.S. Wagner, Insurers as asset managers and systemic risk, 35(12) (2022) 5483-5534.
3. M.F. Aboueleinein, E.A. Hassanein, N. Houssam, R.R. Esily, D.M. Ibrahiem, M. Alnehabi, Z.J.E.S.R. Ahmed, Geopolitical risks and environmental sustainability: Can economic complexity and renewables tip the balance?, 63 (2026) 102027.
4. M.P. Segodi, A.B.J.R. Sibindi, Determinants of life insurance demand: Empirical Evidence from BRICS Countries, 10(4) (2022) 73.
5. J.J.O.E.J. Witkowska, The life and non-life insurance market in the European Union, 18(2) (2023) 157-170.
6. M.E.J.T.R.o.E.S. Yaari, Uncertain lifetime, life insurance, and the theory of the consumer, 32(2) (1965) 137-150.
7. M.F.J.S. Aboueleinein, Optimization, I. Computing, Resilience and Divergence: A Comparative Analysis of Life Insurance Markets in Germany and Japan during the COVID-19 Pandemic and the Interest Rate Trend Reversal (2020-2022), (2026).
8. A.J.Q. Cornaro, Financial resilience of insurance network during Covid-19 pandemic, 57(Suppl 2) (2023) 151-172.
9. J. Hong, S.H.J.R.A. Seog, Health insurance system and resilience to epidemics, 43(1) (2023) 97-114.
10. M.F. Aboueleinein, G.S.S.J.S. Abdalla, Optimization, I. Computing, The Impact of Bancassurance on the Performance and Efficiency of Life Insurance Companies, (2026).
11. S.J.I.E.R. Fischer, A life cycle model of life insurance purchases, (1973) 132-152.
12. R.A.J.T.j.o.F. Campbell, The demand for life insurance: An application of the economics of uncertainty, 35(5) (1980) 1155-1172.
13. G.S.S. Abdalla, M.F. Aboueleinein, H.M.J.S. Noaman, Optimization, I. Computing, Enhancing Fraud Detection in Health Insurance: Deep Neural Network Approaches and Performance Analysis, 14(6) (2025) 3565-3588.
14. M. Koller, Life insurance risk management essentials, Springer Science & Business Media, 2011.
15. J.F.J.R.M. Outreville, I. Review, The relationship between insurance and economic development: 85 empirical papers for a review of the literature, 16(1) (2013) 71-122.
16. T. Beck, I.M. Webb, Determinants of life insurance consumption across countries, World Bank Publications, 2003.
17. C.-H. Chang, C.-C.J.T.G.R. Lee, I. Review, Non-linearity between life insurance and economic development: A revisited approach, 37(2) (2012) 223-257.
18. E. Berdin, H.J.T.G.P.o.R. Gründl, Insurance-Issues, Practice, The effects of a low interest rate environment on life insurers, 40(3) (2015) 385-415.
19. D. Mayers, C.W. Smith Jr, On the corporate demand for insurance, Foundations of insurance economics: Readings in economics and finance, Springer, 1982, pp. 190-205.
20. Y. Croissant, G.J.J.o.s.s. Millo, Panel data econometrics in R: The plm package, 27 (2008) 1-43.
21. J.A.J.E.J.o.t.e.s. Hausman, Specification tests in econometrics, (1978) 1251-1271.
22. M. Arellano, O.J.J.o.e. Bover, Another look at the instrumental variable estimation of error-components models, 68(1) (1995) 29-51.
23. K. A. E. Gad, F. A. Atia, M. Moussa, H. M. Hamouda, M. F. Aboueleinein, The Fractional Exponential Distribution: A Gamma Subfamily from Conformable Calculus, Results in Engineering, 111224 (2026).
24. A. A. Ali, M. Ali, K. A. Gad, S. M. Assar, Some Characterizations Based on Generalized Order Statistics from Power Inverted Topp–Leone Distribution, Advances in Nonlinear Variational Inequalities, 28(6s), 349–365 (2025).
25. A. S. Eldeeb, I. B. Abdul-Moniem, K. A. E. Gad, A New Discrete Power Function Distribution with Its Mathematical Properties, Journal of Information Systems Engineering & Management, 10(44s) (2025).
26. K. A. E. Gad, I. B. Abdul-Moniem, S. M. Assar, Some Characterizations Based on Generalized Order Statistics from Weibull-Weibull Distribution, The Egyptian Statistical Journal, 67(2), 1–16 (2023).
27. A. A. Abdul-Rahman, I. B. Abdul-Moniem, K. A. E. Gad, S. M. Assar, Some Characterizations Based on GOS from Weibull-Weibull Distribution (Part II), Egyptian Statistical Journal, 67(2) (2023).
28. I. B. Abdul-Moniem, K. A. E. Gad, S. M. Assar, Characterization of Phani Distribution Based on Generalized Order Statistics, Sohag Journal of Sciences, 9(1), 32–39 (2024).
29. A. A. Abdul-Rahman, I. B. Abdul-Moniem, K. A. E. Gad, S. M. Assar, Some Characterizations Based on Generalized Order Statistics from Weibull-Family of Life Distributions, Proceedings of the 56th Annual International Conference on Data Sciences (2024).
30. A. S. Hassan, I. B. Abdul-Moniem, K. A. E. Gad, A Generalized Transmuted Moment Exponential Distribution: Properties and Application, Academic Journal of Applied Mathematical Sciences, 6(5), 41–52 (2020).