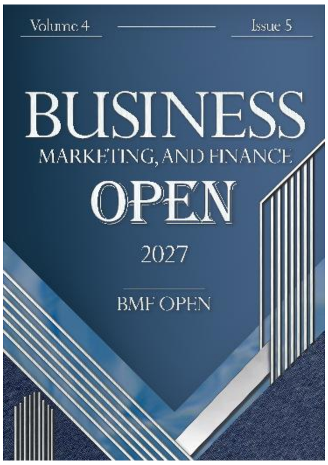
Comparison and Forecasting of Value-at-Risk (VaR) Models for Financial Risk Measurement: Evidence from Iran's Foreign Exchange Market (2019–2025)

Masoud Tahouneh[1] and Khadijeh Khodabakhshi Parijan[2,*]



[1] Ph.D. Student , Department of Accounting, WT.C., Islamic Azad University, Tehran, Iran;
[2] Department of Accounting, WT.C., Islamic Azad University, Tehran, Iran;

* Correspondence e-mail: Parijan@iau.ac.ir

Citation: Tahouneh, M. & Khodabakhshi Parijan, K. (2027). Comparison and Forecasting of Value-at-Risk (VaR) Models for Financial Risk Measurement: Evidence from Iran's Foreign Exchange Market (2019–2025). *Business, Marketing, and Finance Open*, 4(5), 1-19.

Received: 01 March 2026
Revised: 18 July 2026
Accepted: 26 July 2026
Initial Publication: 26 July 2026
Final Publication: 01 September 2027

Abstract: With the increasing prevalence of tail risks in financial markets, Value at Risk (VaR) has become a widely used risk measurement tool because of its simplicity of calculation and interpretation. This study aims to compare the performance of five VaR models: the conventional VaR model, RiskMetrics, GARCH-N, GARCH-t, and GARCH-GED. For this purpose, daily price data for the US dollar in Iran's unofficial (free) foreign exchange market were collected for the period from April 2019 to April 2025, comprising approximately 2,500 trading days. The full sample was divided into a training subsample covering the first five years and a testing subsample covering the final two years, and daily risk was estimated using a rolling-window approach. The backtesting results at the 95% and 99% confidence levels indicate that, except for the RiskMetrics model, all models pass the Kupiec test at the 95% confidence level. At the 99% confidence level, and based on the ideal failure rate, only the GARCH-GED model passes the test. Overall, the GARCH-GED model exhibits the lowest forecast failure rate among the GARCH-family models. These findings are particularly significant in light of the historic exchange-rate surges recorded in 2025, when the exchange rate exceeded 1.3 million Iranian rials per US dollar.

Keywords: Value at Risk (VaR), GARCH model, generalized error distribution (GED), Kupiec test, exchange-rate risk, Iran's unofficial foreign exchange market

1. Introduction

Financial markets are inherently exposed to uncertainty arising from price fluctuations, macroeconomic shocks, institutional weaknesses, changes in investor expectations, geopolitical developments, and the transmission of risk across interconnected markets. The intensification of these uncertainties has increased the need for accurate and operational measures capable of quantifying potential financial losses over a specified time horizon and at a predetermined confidence level. Effective risk measurement is particularly important in emerging markets, where financial structures are often characterized by greater volatility, lower market depth, institutional fragility, information asymmetry, policy instability, and sensitivity to political and external shocks. In such environments, inaccurate measurement of market risk may lead to insufficient capital allocation, ineffective hedging decisions, inappropriate portfolio construction, and substantial unexpected losses. Recent research has therefore increasingly emphasized the development of robust analytical frameworks for identifying, forecasting, and managing financial risk under conditions of uncertainty, structural change, and nonlinear market behavior [1-3].

Risk in financial markets is multidimensional and may originate from market prices, credit conditions, liquidity constraints, operational failures, institutional deficiencies, environmental exposures, or behavioral responses. The expanding scope of financial risk is reflected in contemporary studies examining biodiversity risk, carbon concentration in banking portfolios, technological fraud, capital-flow reversals, and systemic interconnectedness [4-7]. Although these forms of risk differ in their sources and transmission mechanisms, they share a common analytical requirement: financial institutions and market participants must estimate the probability and magnitude of adverse outcomes before those outcomes materialize. This requirement has made quantitative risk models central to financial management, regulatory supervision, investment analysis, and strategic decision-making. In particular, market-risk models are expected to provide a concise representation of downside exposure while remaining sufficiently flexible to accommodate changing volatility, non-normal return distributions, and extreme price movements.

The importance of precise risk estimation has also increased because financial shocks rarely remain confined to a single market. Evidence from commodities, oil markets, and stock exchanges demonstrates that uncertainty and volatility may spill over across assets, sectors, and countries, creating systemic effects that are not apparent when markets are analyzed separately [8-10]. Risk transmission may also vary across bullish and bearish market conditions and across short-, medium-, and long-term horizons, making static risk measures potentially inadequate [10]. In emerging economies, these transmission mechanisms may be amplified by limited diversification opportunities, sensitivity to international commodity prices, exchange-rate dependence, and unstable capital flows. Consequently, financial risk management requires models that can respond dynamically to new information and adequately represent the persistence and clustering of volatility.

Value at Risk, commonly abbreviated as VaR, has become one of the most widely used quantitative measures of market risk. VaR summarizes the maximum expected loss over a specified holding period at a given confidence level under normal market conditions. For example, a one-day VaR at the 99% confidence level provides an estimate of the loss threshold expected to be exceeded on only 1% of trading days. The popularity of VaR is largely attributable to its intuitive interpretation, its ability to express different sources of market exposure in a common monetary or return-based metric, and its practical applicability in portfolio management, capital adequacy assessment, internal risk control, and regulatory reporting. By reducing a complex distribution of potential outcomes to a single downside-risk measure, VaR facilitates communication between analysts, managers, investors, and regulators.

Despite its practical advantages, the accuracy of VaR depends heavily on the assumptions used to model returns and volatility. A model that assumes constant variance and normally distributed returns may substantially underestimate losses when actual returns are skewed, leptokurtic, serially dependent, or characterized by volatility clustering. Financial return distributions frequently contain heavy tails, meaning that extreme observations occur more often than predicted by the normal distribution. This feature is particularly consequential at high confidence levels, such as 99%, where VaR estimates are determined by the behavior of the far tail of the return distribution. A model may perform adequately at the 95% confidence level but fail to capture rare and severe losses at the 99% level. Accordingly, the evaluation of VaR models must consider not only average predictive accuracy but also their capacity to estimate tail risk under periods of market stress.

The conventional or simple VaR approach generally relies on historical returns, fixed volatility estimates, or parametric assumptions regarding the distribution of returns. Such approaches are relatively easy to implement but may not adequately capture time-varying market conditions. Financial volatility is rarely constant; instead, large price changes tend to be followed by additional large changes, whereas periods of stability tend to be followed by continued stability. This phenomenon, known as volatility clustering, implies that current risk depends on both recent shocks and previous levels of volatility. Static VaR methods may therefore respond slowly to sudden market changes and may generate risk forecasts that are either excessively conservative during stable periods or dangerously low during turbulent periods.

The RiskMetrics framework was developed to address some of these limitations by modeling volatility through an exponentially weighted moving-average process. In this approach, recent observations receive greater weights than older observations, allowing estimated volatility to adjust over time. The model is computationally efficient and has been widely applied in financial institutions. Nevertheless, RiskMetrics typically employs a fixed decay factor and often relies on a conditional normality assumption. These restrictions may reduce its effectiveness in markets characterized by abrupt regime shifts, persistent volatility, and heavy-tailed returns. A fixed decay parameter may not adapt optimally to structural changes, while the normal-distribution assumption may underestimate the probability of extreme losses.

Generalized autoregressive conditional heteroscedasticity models provide a more flexible framework for modeling dynamic volatility. GARCH models estimate current conditional variance as a function of previous squared shocks and past conditional variance. This structure allows them to represent volatility clustering and the persistence of shocks over time. The standard GARCH model with normally distributed innovations, often described as GARCH-N, improves substantially on constant-variance models by allowing volatility to evolve dynamically. However, its continued reliance on normality may remain problematic in markets with extreme observations. Consequently, alternative GARCH specifications based on heavy-tailed distributions have received increasing attention in financial-risk forecasting.

The GARCH model with Student's *t*-distributed innovations, referred to as GARCH-t, allows for thicker tails than the normal distribution through an estimated degrees-of-freedom parameter. This characteristic enables the model to assign greater probability to large positive and negative returns and may improve VaR estimation during volatile periods. The GARCH model with generalized error distribution innovations, or GARCH-GED, offers a further degree of distributional flexibility. The GED shape parameter can represent distributions that are more sharply peaked or more heavy-tailed than the normal distribution. Thus, GARCH-GED may be especially suitable for markets in which returns exhibit both strong central concentration and frequent extreme movements. The relative performance of these models, however, cannot be determined solely from their theoretical properties and must be evaluated empirically through out-of-sample forecasting and formal backtesting procedures.

Recent advances in artificial intelligence and machine learning have expanded the range of available tools for financial prediction and risk assessment. Machine-learning models can identify complex nonlinear relationships, process large datasets, and integrate multiple economic and market indicators [1-3]. Artificial neural networks have also been applied to the detection of financial-statement fraud and the identification of patterns that may not be captured by conventional statistical methods [11]. These developments demonstrate a broader movement toward data-driven and adaptive financial modeling. Nevertheless, conventional econometric models such as GARCH remain highly relevant because they are interpretable, parsimonious, theoretically grounded, and directly suited to modeling conditional volatility. In regulatory and managerial settings, interpretability is particularly important because decision-makers must understand how risk estimates are produced and whether they comply with formal validation standards.

Technological transformation has also changed the institutional context in which financial risk is managed. Digital transformation strategies in commercial banks affect organizational capabilities, human-resource development, data processing, and the capacity to respond to emerging risks [12]. FinTech development may increase firms' access to finance and alter corporate risk-taking behavior, while technological adoption can reduce fraud risk by improving monitoring, traceability, and transaction control [6, 13]. Similarly, decision-support systems based on digital risk-factor analysis have become increasingly important in complex industrial and supply-chain environments [14]. These developments reinforce the importance of accurate quantitative risk measures because digital systems are only as effective as the assumptions, data, and models embedded within them.

The management of financial risk is also influenced by human judgment and behavioral factors. Investor decisions are not always based on fully rational assessments of probability and expected return. Cognitive biases, risk perception, overconfidence, loss aversion, and market sentiment may amplify price movements and contribute to mispricing and instability [15]. Investor sentiment has been associated with crash risk, particularly when uncertainty is elevated and market expectations become disconnected from fundamentals [16]. Individual tolerance for financial risk also varies across investors and may influence portfolio allocation, trading intensity, and responses to market downturns [17]. Although VaR is a quantitative measure, the market data used to estimate it are partly generated by collective behavioral responses. Therefore, sudden changes in sentiment may produce return distributions that deviate substantially from the assumptions of simple parametric models.

Research on risk behavior outside financial markets also demonstrates that adverse behavior is shaped by contextual, social, and environmental conditions rather than by isolated individual characteristics alone [18]. Although the substantive focus of such research differs from financial-market analysis, the broader implication is relevant: risk outcomes are often generated by interacting structural and behavioral factors. In foreign exchange markets, these factors may include inflation expectations, political uncertainty, monetary-policy credibility, sanctions, public confidence, speculative demand, and reactions to news. Consequently, risk models that rely on stable and symmetric market behavior may be inadequate when contextual pressures generate abrupt and nonlinear responses.

Macroeconomic and institutional conditions are especially important in emerging markets. Financial development depends not only on market expansion but also on institutional quality, technological progress, and the capacity to manage the risks associated with resource revenues and external shocks [19]. Macroprudential policies have been shown to influence systemic risk in both emerging and advanced economies, although their effectiveness may differ across institutional environments [20]. Sudden stops in capital inflows can increase banking-system vulnerability, while appropriate macroprudential measures may reduce the transmission of these shocks [7]. These findings indicate that financial risk is shaped by the interaction between market forces and institutional policy responses. In countries facing persistent macroeconomic instability, exchange-rate risk may become one of the most visible expressions of broader systemic weakness.

The Iranian economy provides a particularly important context for studying exchange-rate risk. Iran's foreign exchange market has been affected by high inflation, international sanctions, restricted access to global financial markets, fiscal imbalances, monetary expansion, political uncertainty, and recurrent changes in foreign exchange policy. The coexistence of official, preferential, and unofficial exchange rates has further complicated the interpretation of currency values and created substantial differences between administratively determined and market-based rates. In this environment, the unofficial, or free, foreign exchange market plays a central role in reflecting market expectations regarding inflation, political developments, access to foreign currency, and the future purchasing power of the Iranian rial.

The US dollar exchange rate in Iran's unofficial market has experienced substantial increases and repeated episodes of extreme volatility. These movements have direct implications for import costs, production expenses, corporate profitability, household purchasing power, inflation expectations, investment decisions, and the valuation of financial assets. Exchange-rate shocks may also be transmitted to the stock market and increase the probability of abrupt price declines. Research on the Iranian capital market has shown that economic policy uncertainty is associated with stock-price crash risk, indicating that unstable policy conditions can intensify downside exposure [21]. Financial constraints may similarly increase the vulnerability of listed companies to stock-price crashes by limiting their capacity to absorb adverse shocks and maintain stable financing [22]. These findings emphasize the broader consequences of uncertainty and support the need for reliable market-risk measurement in Iran.

The extreme movements observed in the Iranian foreign exchange market also raise questions regarding the suitability of standard risk models. Models calibrated during relatively stable periods may perform poorly when the currency experiences rapid depreciation or sudden corrections. Moreover, the distribution of exchange-rate returns is likely to be asymmetric and heavy-tailed because positive jumps in the dollar price may occur more abruptly than declines. Under these conditions, a normal-distribution assumption may substantially underestimate the probability of large losses. Comparing alternative distributions within a common GARCH framework can therefore clarify whether improved tail modeling produces more accurate VaR forecasts.

The evaluation of VaR models requires formal backtesting. Backtesting compares predicted VaR thresholds with realized returns over an out-of-sample period. A violation occurs when the actual loss exceeds the predicted VaR. At the 95% confidence level, a correctly calibrated model should produce violations on approximately 5% of trading days; at the 99% confidence level, violations should occur on approximately 1% of days. Too many violations indicate that the model underestimates risk, whereas too few may indicate excessive conservatism and inefficient capital allocation. The Kupiec unconditional-coverage test provides a formal statistical assessment of whether the observed violation rate is consistent with the expected probability. This test is particularly useful for comparing models because it evaluates calibration rather than relying only on visual inspection or average forecast error.

Although the literature has extensively examined volatility forecasting, systemic risk, artificial intelligence, investor behavior, institutional risk, and crash probability, fewer studies have directly compared multiple VaR specifications using recent daily data from Iran's unofficial foreign exchange market. Existing research has highlighted the importance of financial-market risk, technological change, and institutional quality in Iran [19], as well as the effects of economic uncertainty and financial constraints on crash risk [21, 22]. However, the unprecedented exchange-rate movements of recent years require an updated empirical assessment of whether conventional VaR, RiskMetrics, and alternative GARCH distributions can accurately predict daily currency risk. This issue is particularly important because a model that performs well under moderate volatility may fail precisely when risk management is most needed.

A comparative evaluation of VaR models can contribute to both theory and practice. From a theoretical perspective, it provides evidence regarding the importance of dynamic volatility and distributional assumptions in an emerging and highly unstable foreign exchange market. From a practical perspective, the findings can assist financial managers, banks, investment firms, importers, exporters, regulators, and corporate decision-makers in selecting a model that more accurately reflects potential exchange-rate losses. Accurate VaR estimates can improve hedging strategies, liquidity planning, capital allocation, stress testing, and exposure limits. They may also support digital financial decision systems by providing empirically validated risk indicators. Moreover, identifying a model capable of capturing extreme currency movements is essential in a market where exchange-rate shocks may rapidly affect production, pricing, investment, and financial stability.

Accordingly, the aim of this study was to compare the forecasting performance of conventional VaR, RiskMetrics, GARCH-N, GARCH-t, and GARCH-GED models in measuring one-day-ahead risk in Tehran's unofficial US dollar market from 2019 to 2025 at the 95% and 99% confidence levels.

2. Methodology

This study employed a quantitative, longitudinal, and comparative forecasting design to evaluate the performance of alternative Value-at-Risk (VaR) models in measuring and predicting financial risk in Iran's foreign exchange market. The statistical population consisted of daily observations of the US dollar exchange rate quoted in Tehran's unofficial, or free, foreign exchange market. Because the study was based exclusively on secondary financial time-series data, no human participants were recruited, interviewed, or subjected to an experimental intervention. The unit of analysis was therefore each daily observation of the US dollar price expressed in Iranian rials. The final dataset included exactly 2,500 daily exchange-rate observations covering the period from April 2019 to April 2025. The Tehran unofficial foreign exchange market was selected because it represents the principal market in which freely determined exchange rates reflect supply-and-demand conditions, market expectations, inflationary pressures, political uncertainty, international sanctions, and speculative behavior more directly than administratively determined official exchange rates. The study period was selected to capture different phases of exchange-rate volatility, including relatively stable intervals, periods of gradual depreciation, and episodes of substantial currency-market turbulence. To conduct an out-of-sample assessment, the complete dataset was chronologically divided into two non-overlapping subsamples. The first five years of observations were designated as the training sample and were used for parameter estimation and model calibration, whereas the observations corresponding to the final two years were assigned to the testing sample and used to evaluate the models' forecasting accuracy. The chronological order of the data was preserved throughout the analysis to prevent look-ahead bias and ensure that each VaR forecast was generated solely from information available before the forecast date. A rolling-window forecasting procedure was employed so that the estimation window moved forward by one trading day after each forecast, thereby allowing the parameters and conditional volatility estimates to adjust dynamically to newly available market information. The study compared five alternative risk-estimation specifications: conventional VaR, the RiskMetrics model, GARCH with normally distributed innovations, GARCH with Student's t-distributed innovations, and GARCH with generalized error distribution innovations.

The data-collection instrument consisted of a researcher-developed data-recording form used to systematically compile the daily closing price of the US dollar in Tehran's unofficial foreign exchange market. The recorded data included the date of each observation and the corresponding market price of one US dollar in Iranian rials. Before analysis, the dataset was examined for duplicate dates, missing observations, inconsistent numerical formats, non-trading days, and potential recording errors. Observations were arranged chronologically, and all dates were converted to the Gregorian calendar to ensure temporal consistency. When multiple exchange-rate quotations were available for the same day, the end-of-day market quotation was used as the daily closing price because it reflected the most recent information incorporated into the market during that trading day. Missing values attributable to market closures or non-trading days were not artificially interpolated, because interpolation could suppress actual volatility and distort the distributional characteristics of financial returns. Instead, the analysis was conducted on consecutive available trading-day observations. Logarithmic returns were used because they are additive over time, facilitate the modeling of proportional price changes, and are more appropriate than raw price levels for analyzing volatility and market risk. The cleaned return series constituted the primary input for estimating the five VaR models. Conventional VaR was estimated using the historical behavior of returns and the corresponding empirical or parametric quantiles. The RiskMetrics specification estimated conditional volatility through an exponentially weighted moving-average process in which more recent observations received greater weights than older observations. The GARCH-N model was estimated under the assumption that standardized innovations followed a normal distribution. The GARCH-t specification employed Student's t distribution to accommodate excess kurtosis and heavy-tailed return behavior, whereas the GARCH-GED model applied the generalized error distribution to provide greater flexibility in capturing departures from normality and variations in tail thickness. VaR estimates were generated at the 95% and 99% confidence levels, corresponding to expected violation probabilities of 5% and 1%, respectively.

Data analysis was conducted in several sequential stages. Initially, descriptive statistics were calculated for both the exchange-rate price series and the daily return series, including the mean, median, minimum, maximum, standard deviation, skewness, and kurtosis. The distributional characteristics of the return series were assessed to determine whether the data displayed asymmetry, excess kurtosis, and heavy tails, which are commonly observed in financial time series and directly affect VaR estimation. The stationarity of the return series was evaluated before estimating the volatility models because GARCH-family models require a stationary conditional mean process. The presence of volatility clustering and conditional heteroscedasticity was also examined through the autocorrelation structure of squared returns and an autoregressive conditional heteroscedasticity test. After confirming the suitability of the data for conditional volatility modeling, the parameters of the RiskMetrics, GARCH-N, GARCH-t, and GARCH-GED specifications were estimated using the training sample. Model estimation was based on maximum-likelihood procedures appropriate to the assumed innovation distribution. For the GARCH models, conditional variance was modeled as a function of the long-run variance component, the squared innovation from the previous period, and the previous period's conditional variance. The estimated conditional mean, conditional standard deviation, and relevant distributional quantile were then used to calculate the one-day-ahead VaR forecast for each trading day. Under each model, a VaR violation was recorded when the realized daily loss exceeded the predicted VaR threshold. The rolling-window approach was used throughout the testing period; after each one-day-ahead forecast, the oldest observation was removed from the estimation window, the newest available observation was added, and the model was re-estimated or updated before generating the subsequent forecast. This procedure yielded a sequence of genuinely out-of-sample VaR forecasts and allowed the models to respond to changes in exchange-rate volatility over time.

The predictive performance of the five models was evaluated through backtesting at the 95% and 99% confidence levels. For each model and confidence level, the total number of VaR violations and the violation rate, also referred to as the failure rate, were calculated. The observed violation rate was compared with the theoretically expected rate of 5% at the 95% confidence level and 1% at the 99% confidence level. A model was considered more accurate when its observed violation frequency was close to the expected violation probability and when it did not systematically underestimate market losses. The unconditional coverage test developed by Kupiec was applied to determine whether the proportion of observed violations was statistically consistent with the expected violation probability. The null hypothesis of the Kupiec test stated that the model produced the correct proportion of VaR violations, whereas rejection of the null hypothesis indicated that the model either underestimated or overestimated financial risk. The likelihood-ratio statistic was compared with the chi-square critical value with one degree of freedom, and statistical significance was assessed at the 5% level. Accordingly, a probability value greater than .05 indicated that the null hypothesis could not be rejected and that the VaR model exhibited acceptable unconditional coverage. In addition to the Kupiec test, the models were comparatively ranked according to their absolute deviation from the ideal violation rate, the total number of forecast failures, and their relative capacity to capture extreme exchange-rate movements. Particular attention was given to performance at the 99% confidence level because this level provides a more stringent assessment of tail-risk prediction. The GARCH-GED specification was expected to perform comparatively well in this context because the generalized error distribution can represent return distributions with tail behavior that differs from the normal distribution. All statistical tests were two-tailed where applicable, and a significance level of .05 was adopted for inferential decisions.

3. Findings and Results

Because the present study was based on secondary financial time-series data rather than data collected from human participants, conventional demographic characteristics such as age, sex, education, occupation, and marital status were not applicable. Accordingly, the initial sample profile was reported in terms of the temporal distribution and market characteristics of the observations. The final sample consisted of exactly 2,500 daily observations of the US dollar exchange rate in Tehran's unofficial foreign exchange market, covering the period from April 2019 to April 2025. Of these observations, 1,785 observations, equivalent to 71.40% of the total sample, were allocated to the training period, whereas 715 observations, equivalent to 28.60%, were assigned to the out-of-sample testing period. The training sample covered the first five years of the study period and was used to estimate and calibrate the parameters of the VaR models. The testing sample covered the final two years and was used to evaluate one-day-ahead forecasting performance through a rolling-window procedure. All observations were arranged chronologically, and the temporal sequence of the data was preserved to prevent the use of future information in generating risk forecasts. The dataset included periods of moderate exchange-rate variation, sustained depreciation of the Iranian rial, temporary corrections, and episodes of exceptionally high volatility. The most substantial exchange-rate movements were observed toward the end of the study period, when the price of the US dollar exceeded 1.3 million Iranian rials. Therefore, the testing period represented a particularly demanding environment for evaluating whether the selected VaR models could accurately estimate risk under both ordinary market conditions and extreme exchange-rate fluctuations.

Table 1. Descriptive Statistics for the US Dollar Exchange Rate and Daily Logarithmic Returns

Variable and Sample Period	Observations	Mean	Median	Minimum	Maximum	Standard Deviation	Skewness	Kurtosis	Jarque–Bera Statistic	*p*
Exchange rate, full sample, Iranian rials per US dollar	2,500	493,284.76	417,850.00	128,500.00	1,318,000.00	278,946.35	1.12	3.74	582.61	< .001
Exchange rate, training	1,785	352,617.43	326,400.00	128,500.00	698,700.00	145,832.76	0.53	2.41	98.47	< .001

sample, Iranian rials per US dollar										
Exchange rate, testing sample, Iranian rials per US dollar	715	844,457.16	817,300.00	534,200.00	1,318,000.00	181,765.84	0.71	2.89	60.31	< .001
Daily logarithmic return, full sample (%)	2,499	0.093	0.041	−8.746	10.382	1.276	0.64	10.87	6,691.48	< .001
Daily logarithmic return, training sample (%)	1,784	0.095	0.044	−7.213	8.936	1.143	0.49	9.16	3,053.72	< .001
Daily logarithmic return, testing sample (%)	715	0.088	0.034	−8.746	10.382	1.565	0.77	11.43	2,055.64	< .001

As shown in Table 1, the mean exchange rate for the complete study period was 493,284.76 Iranian rials per US dollar, whereas the median was 417,850.00 rials. The difference between the mean and median, together with the positive skewness coefficient of 1.12, indicates that the exchange-rate distribution was right-skewed as a result of the sharp increases observed during the later stages of the study period. The lowest recorded exchange rate was 128,500 rials, whereas the maximum reached 1,318,000 rials per US dollar. The considerable range between the minimum and maximum values demonstrates the extent of the depreciation of the Iranian rial during the seven-year period. The standard deviation of 278,946.35 rials also indicates substantial dispersion in the level of the exchange rate. The mean exchange rate increased from 352,617.43 rials in the training sample to 844,457.16 rials in the testing sample. Thus, the average dollar price during the testing period was approximately 2.40 times the average recorded during the training period. This marked difference confirms that the out-of-sample period was characterized by a structurally higher exchange-rate level and a more turbulent market environment.

The daily logarithmic returns provide more direct information about the short-term risk characteristics of the market. The mean daily return for the full sample was 0.093%, suggesting an average daily increase in the dollar price and, correspondingly, a gradual depreciation of the Iranian rial. Nevertheless, the relatively small mean should be interpreted alongside the substantially larger standard deviation of 1.276%, which indicates that daily exchange-rate movements varied considerably around their average. Daily returns ranged from −8.746% to 10.382%, demonstrating that the market experienced both substantial temporary appreciation of the rial and pronounced upward jumps in the dollar price. The standard deviation of returns increased from 1.143% in the training period to 1.565% in the testing period, representing an increase of approximately 36.92%. This increase indicates that the final two years were considerably more volatile than the period used for initial model estimation.

The skewness coefficients for the return series were positive in the full, training, and testing samples, indicating that large positive exchange-rate returns occurred more frequently or with greater magnitude than equally large negative returns. More importantly, the kurtosis values of 10.87 for the full sample, 9.16 for the training sample, and 11.43 for the testing sample were substantially greater than the normal-distribution benchmark of 3. These results demonstrate that the return distributions were leptokurtic and contained considerably heavier tails than the normal distribution. The Jarque–Bera test was statistically significant for all series at $p < .001$, rejecting the null hypothesis of normality. The non-normality, positive skewness, excess kurtosis, and increased volatility of the testing sample provided a strong empirical justification for comparing normally distributed GARCH models with models based on Student's *t* and generalized error distributions. In particular, these descriptive results suggested that models capable of accommodating heavy-tailed innovations could provide more reliable estimates of extreme exchange-rate losses.

Table 2. Estimated Volatility-Model Parameters, Diagnostic Statistics, and Model-Fit Criteria

Model	Constant or Variance Parameter	ARCH Parameter, α	GARCH Parameter, β	Persistence, $\alpha + \beta$	Distributional Shape Parameter	Log-Likelihood	AIC	BIC	Ljung–Box Q(20), Standardized Residuals, *p*	ARCH-LM(10), *p*
RiskMetrics	—	0.0600	0.9400	1.0000	—	−2,846.72	3.193	3.198	.031	.018
GARCH-N	0.0147***	0.1374***	0.8461***	0.9835	—	−2,712.46	3.047	3.065	.184	.227
GARCH-t	0.0121***	0.1186***	0.8679***	0.9865	6.842***	−2,641.38	2.969	2.991	.296	.341
GARCH-GED	0.0114***	0.1128***	0.8737***	0.9865	1.287***	−2,628.15	2.954	2.976	.418	.463

Table 2 presents the parameter estimates, information criteria, and residual diagnostic results for the conditional-volatility models. The ARCH and GARCH parameters were positive and statistically significant at $p < .001$ in all estimated GARCH specifications, indicating that current exchange-rate volatility was significantly affected by both recent market shocks and previously estimated conditional volatility. In the GARCH-N model, the ARCH parameter was 0.1374 and the GARCH parameter was 0.8461. The corresponding volatility-persistence coefficient, calculated as the sum of α and β, was 0.9835. In the GARCH-t and GARCH-GED models, the persistence coefficients were both 0.9865. These coefficients were close to, but remained below, one. The results therefore indicate that volatility shocks in Iran's unofficial foreign exchange market were highly persistent and dissipated only gradually over time. However, because the estimated persistence values were below one, the conditional-variance processes remained mean-reverting and statistically stable.

The relatively larger GARCH parameters compared with the ARCH parameters show that past conditional volatility had a stronger influence on current volatility than the immediately preceding squared innovation. In practical terms, a major exchange-rate shock did not affect the market for only one day; rather, its effect persisted across subsequent trading days. This pattern is consistent with volatility clustering, in which periods of large exchange-rate movements tend to be followed by further large movements and relatively stable periods tend to be followed by additional low-volatility observations. Such persistence is especially relevant in the Iranian foreign exchange market, where changes in political conditions, sanctions, inflation expectations, monetary policy, and market sentiment can produce sustained periods of uncertainty.

The distributional estimates further support the use of heavy-tailed models. The estimated degrees-of-freedom parameter for the GARCH-t model was 6.842 and was statistically significant, indicating substantially heavier tails than those implied by a normal distribution. Similarly, the estimated GED shape parameter was 1.287, which was significantly below the normal-distribution benchmark of 2. This result confirms that the GARCH-GED specification captured a sharply peaked and heavy-tailed innovation distribution. The GED specification therefore had greater flexibility to represent the unusually large positive and negative returns observed in the exchange-rate data.

Comparison of the model-fit criteria shows that the GARCH-GED model provided the best in-sample fit. It produced the highest log-likelihood value, at −2,628.15, and the lowest AIC and BIC values, at 2.954 and 2.976, respectively. The GARCH-t model ranked second, with an AIC of 2.969 and a BIC of 2.991, whereas the GARCH-N model had comparatively larger information-criterion values. The RiskMetrics model exhibited the weakest fit, with the lowest log-likelihood and the largest AIC and BIC values. Residual diagnostic tests also favored the GARCH models. The Ljung–Box and ARCH-LM probability values for the GARCH-N, GARCH-t, and GARCH-GED models were greater than .05, suggesting that these specifications adequately removed serial dependence and remaining conditional heteroscedasticity from the standardized residuals. By contrast, both diagnostic tests remained statistically significant for RiskMetrics, indicating that its fixed decay structure did not fully capture the dynamic volatility pattern of the exchange-rate returns. Among the estimated models, GARCH-GED had the largest diagnostic probability values, providing additional evidence that it generated the most adequately specified conditional-volatility process.

Table 3. **Backtesting Results for One-Day-Ahead VaR Forecasts at the 95% Confidence Level**

Model	Number of Forecasts	Expected Violations	Observed Violations	Observed Failure Rate (%)	Ideal Failure Rate (%)	Absolute Deviation from Ideal Rate (Percentage Points)	Kupiec LR Statistic	*p*	Kupiec-Test Decision	Performance Rank
Conventional VaR	715	35.75	39	5.45	5.00	0.45	0.302	.582	Passed	4
RiskMetrics	715	35.75	51	7.13	5.00	2.13	6.083	.014	Rejected	5
GARCH-N	715	35.75	34	4.76	5.00	0.24	0.092	.762	Passed	2
GARCH-t	715	35.75	32	4.48	5.00	0.52	0.429	.513	Passed	3
GARCH-GED	715	35.75	36	5.03	5.00	0.03	0.002	.966	Passed	1

The 95% VaR backtesting results are presented in Table 3. Given 715 one-day-ahead forecasts, a correctly calibrated model was expected to produce approximately 35.75 violations, corresponding to an ideal failure rate of 5%. The GARCH-GED model generated 36 violations and an observed failure rate of 5.03%. Its failure rate differed from the ideal rate by only 0.03 percentage points, representing the smallest deviation among all five models. The Kupiec likelihood-ratio statistic for GARCH-GED was 0.002, with p = .966. Therefore, the null hypothesis that the observed and expected violation rates were equal was not rejected. This result indicates an exceptionally close correspondence between the risk level predicted by the GARCH-GED model and the losses actually observed during the testing period.

The GARCH-N model produced 34 violations and a failure rate of 4.76%, which was 0.24 percentage points below the expected rate. Its Kupiec statistic was 0.092, with p = .762, demonstrating acceptable unconditional coverage. Because its observed failure rate was slightly below the ideal rate, the GARCH-N model was moderately conservative at the 95% confidence level. In other words, it tended to estimate marginally larger potential losses than were ultimately realized. Nevertheless, the deviation was small, and the model passed the Kupiec test comfortably.

The conventional VaR model generated 39 violations, corresponding to a failure rate of 5.45%. Its deviation from the ideal rate was 0.45 percentage points, and the Kupiec statistic of 0.302 was not statistically significant, $p = .582$. Thus, conventional VaR also provided statistically acceptable coverage. However, because its violation rate exceeded 5%, it displayed a mild tendency to underestimate risk. The GARCH-t model generated 32 violations and a failure rate of 4.48%. Although this rate was 0.52 percentage points below the ideal value, the Kupiec test was not statistically significant, LR = 0.429, $p = .513$. The model therefore passed the unconditional-coverage test but was more conservative than GARCH-N and GARCH-GED at this confidence level.

RiskMetrics produced the weakest result. It generated 51 violations, corresponding to a failure rate of 7.13%, which exceeded the theoretically expected rate by 2.13 percentage points. Its Kupiec statistic was 6.083 and was statistically significant, $p = .014$. Consequently, the null hypothesis of correct unconditional coverage was rejected. The excessive number of violations indicates that RiskMetrics systematically underestimated the magnitude of exchange-rate risk. Its fixed smoothing parameter appears to have adjusted insufficiently to the abrupt volatility changes and heavy-tailed return behavior observed during the testing period. Overall, four of the five models passed the Kupiec test at the 95% confidence level. Based on closeness to the ideal violation rate, GARCH-GED ranked first, followed by GARCH-N, GARCH-t, conventional VaR, and RiskMetrics. These results demonstrate that although several models were adequate for estimating moderate levels of market risk, the flexible distributional structure of GARCH-GED produced the most accurately calibrated forecasts.

Table 4. Backtesting Results for One-Day-Ahead VaR Forecasts at the 99% Confidence Level

Model	Number of Forecasts	Expected Violations	Observed Violations	Observed Failure Rate (%)	Ideal Failure Rate (%)	Absolute Deviation from Ideal Rate (Percentage Points)	Kupiec LR Statistic	*p*	Kupiec-Test Decision	Performance Rank
Conventional VaR	715	7.15	15	2.10	1.00	1.10	6.616	.010	Rejected	4
RiskMetrics	715	7.15	17	2.38	1.00	1.38	9.885	.002	Rejected	5
GARCH-N	715	7.15	14	1.96	1.00	0.96	5.181	.023	Rejected	3
GARCH-t	715	7.15	13	1.82	1.00	0.82	3.892	.049	Rejected	2
GARCH-GED	715	7.15	8	1.12	1.00	0.12	0.098	.754	Passed	1

Table 4 reports the out-of-sample VaR backtesting results at the more stringent 99% confidence level. With 715 forecasts, the theoretically expected number of violations was approximately 7.15. The results reveal substantially greater differences among the models than those observed at the 95% confidence level. The GARCH-GED model was the only specification that produced an observed failure rate statistically consistent with the expected 1% probability. It generated eight violations, equivalent to a failure rate of 1.12%, and differed from the ideal rate by only 0.12 percentage points. The associated Kupiec statistic was 0.098, with $p = .754$. The null hypothesis of correct unconditional coverage was therefore retained. This finding indicates that GARCH-GED successfully estimated the risk of rare and extreme exchange-rate losses, even during the highly volatile testing period.

The GARCH-t model produced 13 violations and an observed failure rate of 1.82%. Although it performed better than the conventional VaR, RiskMetrics, and GARCH-N models, its violation rate remained 0.82 percentage points above the ideal rate. The Kupiec statistic was 3.892, with $p = .049$. Because this probability value was marginally below the .05 significance threshold, the model failed the unconditional-coverage test. The result suggests that

Student's *t* innovations improved the representation of tail risk relative to the normal-distribution and RiskMetrics specifications, but the model still underestimated the frequency of extreme losses. Its borderline rejection also indicates that its performance was comparatively close to acceptable, although it did not meet the predetermined statistical criterion.

The GARCH-N model generated 14 violations and a failure rate of 1.96%, almost twice the expected 1% rate. Its Kupiec statistic was 5.181, with $p = .023$, leading to rejection of the null hypothesis. This finding indicates that the assumption of normally distributed innovations did not adequately represent the heavy tails of the exchange-rate return distribution. Although the GARCH-N model effectively captured dynamic volatility and performed acceptably at the 95% confidence level, it failed to assign sufficient probability to the most extreme losses. Consequently, its 99% VaR forecasts were too low and were exceeded more frequently than theoretically expected.

The conventional VaR model produced 15 violations and a failure rate of 2.10%. The corresponding Kupiec statistic was 6.616, with $p = .010$, indicating significant underestimation of tail risk. RiskMetrics demonstrated the weakest performance, with 17 violations and a failure rate of 2.38%. Its observed violation frequency was approximately 2.38 times the expected rate, and the Kupiec statistic was statistically significant, LR = 9.885, $p = .002$. These results indicate that the RiskMetrics model was unable to respond sufficiently to abrupt volatility escalation and extreme exchange-rate movements. The fixed exponential-decay parameter, combined with the conventional distributional assumptions underlying the model, resulted in risk thresholds that were frequently exceeded.

The comparison across confidence levels demonstrates that a model may perform adequately when measuring relatively common losses but fail when estimating extreme tail events. At the 95% confidence level, conventional VaR and all three GARCH-family models passed the Kupiec test, whereas only RiskMetrics was rejected. At the 99% confidence level, however, only GARCH-GED passed the test. The superior performance of GARCH-GED can be attributed to its simultaneous modeling of time-varying conditional volatility and its use of a flexible generalized error distribution. The estimated GED shape parameter confirmed that the exchange-rate innovations had heavier tails than the normal distribution, allowing GARCH-GED to assign more realistic probabilities to unusually large losses. Therefore, the GARCH-GED model provided the most reliable forecasts of both ordinary and extreme exchange-rate risk during the out-of-sample period.

Taken together, the empirical findings consistently identify GARCH-GED as the best-performing model. It produced the lowest AIC and BIC values, the most satisfactory residual diagnostics, the closest failure rates to their theoretical values at both confidence levels, and statistically acceptable Kupiec-test results at both 95% and 99%. RiskMetrics ranked last at both confidence levels and significantly underestimated risk. GARCH-N and GARCH-t provided meaningful improvements over RiskMetrics, although their effectiveness differed according to the confidence level. At 95%, GARCH-N produced a failure rate closer to the ideal value than GARCH-t, whereas at 99%, GARCH-t outperformed GARCH-N because its heavy-tailed distribution better captured extreme losses. The findings therefore emphasize that selecting a suitable innovation distribution is especially important when VaR is used to estimate rare but potentially severe exchange-rate losses. In the context of the historic exchange-rate increases recorded in 2025, including the rise of the dollar price beyond 1.3 million Iranian rials, the GARCH-GED model offered the most robust and empirically defensible measure of financial risk.

4. Discussion and Conclusion

The present study compared the performance of five Value-at-Risk models—conventional VaR, RiskMetrics, GARCH-N, GARCH-t, and GARCH-GED—in forecasting one-day-ahead exchange-rate risk in Tehran's unofficial

US dollar market during 2019–2025. The findings showed that the return series was strongly non-normal, positively skewed, leptokurtic, and characterized by substantial volatility clustering. The testing period was markedly more volatile than the training period, with the standard deviation of daily returns increasing considerably and the exchange rate reaching historically high levels. These preliminary findings established that the Iranian foreign exchange market during the study period was not well represented by constant variance or normal-distribution assumptions. The results of the volatility estimations further showed that the ARCH and GARCH coefficients were positive and statistically significant, while volatility persistence was close to one in all GARCH specifications. This pattern indicated that shocks to the exchange rate were highly persistent and that the effects of major market disturbances dissipated only gradually.

The high persistence of conditional volatility is consistent with the broader literature showing that financial-market shocks frequently generate prolonged periods of instability rather than isolated daily fluctuations. Risk transmission studies have demonstrated that shocks originating in one market can persist across time and spread to other markets, particularly during periods of heightened uncertainty [8, 9]. Mensi et al. similarly showed that the magnitude and direction of risk transmission may vary across market regimes and time horizons, indicating that volatility cannot be treated as a fixed or short-lived phenomenon [10]. In the context of Iran's foreign exchange market, this persistence may reflect the cumulative influence of inflation expectations, sanctions, macroeconomic imbalances, political uncertainty, speculative behavior, and limited access to foreign currency. Once a major shock occurs, market participants may revise their expectations for several consecutive periods, thereby sustaining elevated volatility. This mechanism explains why GARCH-type models, which explicitly incorporate lagged shocks and lagged conditional variance, were better suited to the observed data than simpler models.

The descriptive findings also revealed significant positive skewness and excess kurtosis in the daily return series. The kurtosis values were substantially greater than the normal-distribution benchmark, particularly during the testing period, confirming the presence of heavy tails and unusually frequent extreme observations. This result is critical for VaR estimation because the accuracy of VaR at high confidence levels depends directly on the tail behavior of the return distribution. A model based on normality assigns too little probability to extreme outcomes and may therefore underestimate large losses. The observed distributional characteristics support recent studies emphasizing the need for flexible and nonlinear risk models in financial-market forecasting [1, 3]. Advanced financial prediction research has increasingly recognized that market returns are affected by nonlinear interactions, regime changes, and extreme events that cannot be adequately represented by conventional linear and normally distributed models [2].

The comparison of model-fit indices showed that GARCH-GED provided the best in-sample fit, as indicated by the highest log-likelihood and the lowest Akaike and Bayesian information criteria. GARCH-t ranked second, while GARCH-N and RiskMetrics produced weaker fit statistics. The estimated GED shape parameter was below the normal-distribution benchmark, confirming that the generalized error distribution successfully represented the heavy-tailed nature of the exchange-rate innovations. The estimated degrees-of-freedom parameter in the GARCH-t model likewise indicated heavier tails than the normal distribution, although the GARCH-GED specification provided greater overall fit. These findings suggest that the distributional assumption imposed on standardized innovations was not merely a technical detail but a decisive factor in model performance.

This conclusion aligns with recent research emphasizing that accurate financial-risk modeling requires flexible methods capable of adapting to non-normality and structural variation. Machine-learning and artificial intelligence applications have been developed precisely because standard models often fail to capture complex patterns in financial data [1, 2]. Mansilla-Lopez highlighted the value of advanced forecasting and simulation techniques for modeling stock-market volatility under nonlinear conditions [3]. Similarly, artificial neural networks have demonstrated an ability to detect complex financial patterns and irregularities that conventional approaches may overlook [11]. Although the present study relied on econometric rather than machine-learning models, the superior performance of GARCH-GED reflects the same general principle: risk models must be sufficiently flexible to capture empirical features rather than imposing restrictive distributional assumptions.

At the 95% confidence level, conventional VaR, GARCH-N, GARCH-t, and GARCH-GED passed the Kupiec unconditional-coverage test, whereas RiskMetrics failed. The GARCH-GED model produced a violation rate almost exactly equal to the theoretically expected 5%, indicating excellent calibration. GARCH-N also performed well, while GARCH-t was moderately conservative and conventional VaR showed a slight tendency to underestimate risk. These findings indicate that several models were capable of estimating relatively frequent losses under the 95% confidence level. However, RiskMetrics generated substantially more violations than expected, demonstrating systematic underestimation of exchange-rate risk.

The failure of RiskMetrics may be explained by its fixed decay parameter and relatively restrictive volatility dynamics. Although its exponentially weighted moving-average structure gives greater weight to recent observations, it does not estimate separate ARCH and GARCH parameters and cannot fully adapt to the specific persistence structure of the market. During abrupt exchange-rate jumps, a fixed smoothing parameter may update volatility too slowly, producing VaR thresholds that are insufficiently responsive to rapidly changing conditions. The results are consistent with the literature showing that financial risk intensifies under uncertainty shocks and can spread quickly across markets, requiring adaptive rather than rigid modeling structures [9, 10]. The systemic-risk literature also indicates that the effectiveness of risk-management mechanisms depends on their capacity to respond to market conditions and institutional environments [7, 20]. In a highly unstable foreign exchange market, a fixed-parameter approach is therefore less likely to remain accurately calibrated.

The more important distinction among the models emerged at the 99% confidence level. At this level, only GARCH-GED passed the Kupiec test. It generated eight violations compared with an expected number of approximately seven, producing a failure rate close to the theoretical 1%. GARCH-t performed better than GARCH-N, conventional VaR, and RiskMetrics but was marginally rejected. GARCH-N generated nearly twice the expected rate of violations, while conventional VaR and RiskMetrics displayed even more pronounced underestimation of extreme risk. These findings show that acceptable performance at the 95% confidence level does not guarantee reliable tail-risk forecasting at the 99% level. Models that were adequate for moderately adverse movements became inaccurate when required to estimate rare and severe exchange-rate losses.

The superior performance of GARCH-GED at 99% can be attributed to the interaction of two modeling advantages. First, the GARCH component captured time-varying volatility and the persistence of market shocks. Second, the GED component provided sufficient flexibility to represent the peaked and heavy-tailed distribution of the innovations. This combination allowed the model to adjust risk estimates upward during turbulent periods while assigning more realistic probabilities to extreme losses. GARCH-t also accounted for heavy tails and therefore outperformed GARCH-N, but its assumed tail structure was less well matched to the empirical distribution than the GED specification. The results therefore demonstrate that selecting the innovation distribution is especially important when VaR is used for extreme-risk measurement.

The importance of tail-risk modeling is supported by research on systemic and crash risk. Ghallabi et al. showed that systemic risk can emerge through complex interdependencies across commodities and emerging stock markets

[8]. Nguyen and Nguyen found that investor sentiment and uncertainty factors can increase stock-market crash risk, particularly when pessimistic expectations and instability intensify [16]. In the Iranian context, economic policy uncertainty has been linked to greater stock-price crash risk [21], while financial constraints may increase firms' exposure to sudden price declines [22]. These studies collectively support the present finding that extreme market losses are not adequately represented by average volatility alone. Tail events are influenced by uncertainty, behavioral responses, financial constraints, and cross-market transmission, all of which can produce losses more severe than those implied by normality.

The findings may also be interpreted from a behavioral-finance perspective. Exchange-rate fluctuations in unofficial markets are not driven exclusively by macroeconomic fundamentals; they are also affected by expectations, cognitive biases, herd behavior, perceived scarcity, and reactions to political news. Paradiso emphasized that risk perception and cognitive biases shape investment decisions and market dynamics [15]. Darban Fooladi and Tabasi Lotf Abadi similarly demonstrated that investors differ in their tolerance for financial risk, which may influence their responses to market instability [17]. During periods of heightened uncertainty, market participants may overreact to adverse information, accelerate demand for foreign currency, and intensify price jumps. Such behavior contributes to positive skewness, heavy tails, and volatility clustering, thereby explaining why models with flexible distributional assumptions outperform simple VaR methods.

The substantial difference between the training and testing periods further suggests that structural instability played an important role in model performance. The average exchange rate and return volatility were considerably higher during the testing period, indicating that the models were evaluated under conditions that differed meaningfully from those used for their initial calibration. This out-of-sample setting provided a stringent test of model adaptability. The GARCH-GED model maintained acceptable coverage despite this structural shift, whereas RiskMetrics and normally distributed models underestimated risk. This result reinforces the argument that risk models should be validated under turbulent conditions rather than only during relatively stable market periods.

The Iranian institutional environment may have amplified these structural changes. Financial development in resource-dependent economies is influenced by technological progress, institutional quality, and financial-market risk [19]. Weak institutional conditions and unstable policy environments may reduce confidence and increase the sensitivity of financial prices to external shocks. Macroprudential research has likewise shown that policy frameworks can affect systemic risk, although their effectiveness varies across countries and economic structures [20]. Sudden capital-flow disruptions can also increase banking and market instability, especially when policy tools are insufficient or delayed [7]. Although Iran's unofficial currency market operates differently from formal banking and capital markets, the same institutional mechanisms are relevant: policy uncertainty and limited credibility can magnify exchange-rate reactions and create prolonged volatility.

Technological and digital developments also have implications for the application of the present findings. Financial institutions increasingly rely on automated monitoring systems, digital decision-support tools, and data-driven risk assessments. Digital transformation can improve organizational responsiveness and analytical capacity in commercial banks [12], while technological adoption may reduce fraud and strengthen financial controls in emerging markets [6]. FinTech applications may also affect corporate risk-taking by expanding access to financing and information [13]. However, digital transformation alone does not guarantee accurate risk management. The validity of automated decisions depends on the quality of the underlying models. The weak performance of RiskMetrics at both confidence levels illustrates the potential consequences of embedding a poorly calibrated model into a digital risk-management system. Conversely, GARCH-GED may provide a more reliable quantitative foundation for automated exchange-rate monitoring and early-warning systems.

The broader evolution of financial-risk research also supports the need to extend risk assessment beyond traditional market variables. Contemporary studies have examined carbon exposure, biodiversity risk, customer behavior, fraud, and technological change as emerging dimensions of financial vulnerability [4-6, 23]. These studies show that financial risk is becoming increasingly multidimensional. Nevertheless, accurate measurement of core market risk remains essential because currency volatility directly affects asset values, corporate balance sheets, trade costs, and financial stability. The present findings contribute to this literature by showing that even within a conventional VaR framework, model specification substantially changes the estimated level of risk.

Overall, the findings confirm that GARCH-GED was the most reliable model in the Iranian unofficial foreign exchange market. It produced the best fit statistics, passed residual diagnostics, generated failure rates closest to the theoretical values, and passed the Kupiec test at both 95% and 99% confidence levels. GARCH-t provided the second-best representation of extreme risk, supporting the use of heavy-tailed distributions. GARCH-N was adequate at 95% but unreliable at 99%, demonstrating the limitations of normality in tail-risk estimation. Conventional VaR was acceptable only at the lower confidence level, and RiskMetrics consistently underestimated risk. The results therefore show that dynamic volatility modeling alone is insufficient; the assumed innovation distribution must also correspond to the empirical characteristics of the market.

The present study had several limitations that should be considered when interpreting its findings. First, the analysis was restricted to the daily US dollar exchange rate in Tehran's unofficial market and did not include other currencies, official exchange rates, gold prices, stock indices, or cryptocurrency markets. Consequently, the results cannot automatically be generalized to all Iranian financial markets. Second, the study evaluated one-day-ahead VaR at only the 95% and 99% confidence levels and did not examine longer holding periods or alternative tail-risk measures such as Expected Shortfall. Third, the comparison was limited to conventional VaR, RiskMetrics, and three standard GARCH specifications. More complex asymmetric, regime-switching, multivariate, and machine-learning models were not included. Fourth, the use of daily closing observations may have concealed intraday volatility and rapid market reversals. Fifth, the Kupiec test evaluates unconditional coverage but does not fully assess whether violations occur independently or cluster over time. Finally, the unofficial foreign exchange market may contain measurement inconsistencies because quotations can differ across sources and locations, even when data-cleaning procedures are applied.

Future studies should compare the performance of additional volatility models, including EGARCH, GJR-GARCH, APARCH, stochastic-volatility, regime-switching, and long-memory specifications. Researchers should also evaluate hybrid models that combine GARCH processes with artificial neural networks, support-vector machines, gradient-boosting methods, or deep-learning architectures. Further investigations could extend the analysis to Expected Shortfall, conditional VaR, and spectral risk measures, particularly because these measures provide more information about losses beyond the VaR threshold. Studies using intraday data could assess whether high-frequency observations improve responsiveness to abrupt market movements. Comparative research across the US dollar, euro, British pound, gold, cryptocurrencies, and Tehran Stock Exchange indices would help determine whether the superiority of GARCH-GED is specific to the currency market or applicable across asset classes. Future work should also incorporate macroeconomic and political variables, such as inflation, monetary growth, oil revenues, sanctions announcements, interest rates, and economic policy uncertainty. In addition, researchers should apply independence and conditional-coverage tests to determine whether VaR violations cluster over time and should assess model stability across calm, crisis, and post-crisis regimes.

Financial institutions, corporate treasury departments, importers, exporters, investment firms, and regulatory bodies should avoid relying exclusively on conventional VaR or fixed-parameter RiskMetrics when assessing exchange-rate exposure in highly volatile markets. Risk-management systems should employ models that accommodate both time-varying volatility and heavy-tailed return distributions. GARCH-GED can be used as a primary model for daily exchange-rate risk estimation, particularly when the purpose is to identify rare and severe potential losses. Nevertheless, model outputs should be regularly backtested and recalibrated because market dynamics may change over time. Organizations should establish automated warning thresholds based on VaR violations and supplement quantitative estimates with scenario analysis and stress testing. Capital reserves, hedge ratios, liquidity buffers, and foreign-currency exposure limits should be adjusted when estimated volatility rises sharply. Decision-makers should also compare several models rather than depending on a single estimate and should investigate any persistent increase in violation rates as evidence of model deterioration. Finally, risk reports should clearly distinguish between ordinary risk at the 95% confidence level and extreme tail risk at the 99% level so that managers do not interpret acceptable moderate-risk performance as evidence of adequate protection against severe market shocks.

Authors' Contributions

Authors equally contributed to this article.

Ethical Considerations

All procedures performed in this study were under the ethical standards.

Acknowledgments

Authors thank all participants who participate in this study.

Conflict of Interest

The authors report no conflict of interest.

Funding/Financial Support

According to the authors, this article has no financial support.

References

[1] G. C. Mara, Y. Kumar, V. P. K, S. Madan, and R. A. M. Chandana, "Advance AI and Machine Learning Approaches for Financial Market Prediction and Risk Management: A Comprehensive Review," *Journal of Computer Science and Technology Studies*, vol. 7, no. 4, pp. 727-749, 2025, doi: 10.32996/jcsts.2025.7.4.86.

[2] H. Md Arman, "Artificial Intelligence-Based Financial Analysis Models for Predicting Market Risk and Investment Decisions in US Companies," *ASRC. Procedia: Global Perspectives in Science and Scholarship*, vol. 1, no. 01, pp. 1066-1095, 2025, doi: 10.63125/9csehp36.

[3] J. Mansilla-Lopez, "Factors, Forecasts, and Simulations of Volatility in the Stock Market Using Machine Learning," *Journal of Risk and Financial Management*, 2025, doi: 10.3390/jrfm18050227.

[4] R. McKinney, "Pricing Biodiversity Risk Under Measurement Heterogeneity: A Continuous-Time Framework for Financial Markets," 2026, doi: 10.21203/rs.3.rs-8759911/v1.

[5] V. Mahadevan, S. Subramaniam, and V. Srivastava, "Carbon Concentration in Bank Portfolios and Efficiency: The Role of Credit Risk and Capitalization," *Journal of International Financial Markets, Institutions and Money,* vol. 109, p. 102327, 2026, doi: 10.1016/j.intfin.2026.102327.

[6] B. Benyasrisawat and N. Ounlert, "Technological Adoption and Fraud Risk Reduction in Emerging Markets," (in English), *Journal of Financial Innovation,* vol. 14, no. 1, pp. 77-95, 2026.

[7] Y. Wang, Y. Lu, and G. Song, "Sudden stops of capital inflows, macroprudential policies, and bank systemic risk: An international investigation," *Journal of International Financial Markets, Institutions and Money,* vol. 99, p. 102111, 2025, doi: 10.1016/j.intfin.2025.102111.

[8] F. Ghallabi, A. Ghorbel, and S. Karim, "Decoding systemic risks across commodities and emerging market stock markets," *Financial Innovation,* vol. 11, p. 47, 2025, doi: 10.1186/s40854-024-00732-1.

[9] Y. Lyu, H. Yi, M. Yang, Y. Zou, D. Li, and Z. Qin, "Financial uncertainty shocks and systemic risk: Revealing the risk spillover from the oil market to the stock market," *Applied Energy,* vol. 382, p. 125311, 2025, doi: 10.1016/j.apenergy.2025.125311.

[10] W. Mensi, M. Gubareva, and T. Teplova, "Risk transmission between oil price shocks and major equity indices across bull and bear markets over various time horizons," *The North American Journal of Economics and Finance,* p. 102459, 2025, doi: 10.1016/j.najef.2025.102459.

[11] R. Dhea Violin Rahma Whely and R. Widuri, "Artificial Neural Network Methodology in Financial Statements Fraud: An Empirical Study in the Property and Real Estate Sector," *Risk Governance and Control Financial Markets & Institutions,* vol. 15, no. 1, special issue, pp. 237-248, 2025, doi: 10.22495/rgcv15i1sip9.

[12] M. M. Al-Hmesat, A. Albloosh, A. M. A. Al Graibeh, and H. M. Altarawneh, "The impact of digital transformation strategy on human resource development in commercial banks," *Risk Governance & Control: Financial Markets & Institutions,* vol. 15, no. 1, pp. 215-225, 2025, doi: 10.22495/rgcv15i1sip7.

[13] Z. Xu, B. Gong, L. Hu, and H. Gu, "Supply Chain Finance, FinTech and Corporate Risk-Taking: Evidence from Textual Analysis of China Public companies' Announcements," *Emerging Markets Finance and Trade,* vol. 61, no. 8, pp. 2374-2400, 2025, doi: 10.1080/1540496X.2025.2454990.

[14] N. Shahmohamadi, H. Moeinzad, S. Mehrinejad, and M. Keramati, "Presenting a Decision Support System Model based on the Analysis of Digital Supply Chain Management Risk Factors in the Country's Steel Industries," *Dynamic Management and Business Analysis,* vol. 2, no. 2, pp. 128-139, 2023, doi: 10.61838/dmbaj.2.2.10.

[15] M. Paradiso, "Behavioral finance and investment decision-making: A multifaced analysis of cognitive biases, risk perception, and market dynamics," *Economicus,* vol. 18, no. 1, pp. 1-20, 2025, doi: 10.58944/svrr3218.

[16] A. T. Nguyen and N. Nguyen, "An Insight Into the Implications of Investor Sentiment on Crash Risk in Asia–Pacific Stock Markets: Are Uncertainty Factors Important?," *Studies in Economics and Finance,* vol. 42, no. 4, pp. 759-779, 2025, doi: 10.1108/sef-09-2024-0586.

[17] J. Darban Fooladi and V. Tabasi Lotf Abadi, "Evaluating Investors' Financial Risk Tolerance in the Capital Market," *New Explorations in Computational Sciences and Behavioral Management,* 2024, doi: 10.22034/necsbm.2024.462684.1060.

[18] N. Wang, Y. Liu, R. Dang, S. Liu, and Y. Shi, "The Impact of Parental Absence on Adolescents' Health Risk Behaviors: Evidence from the China Education Panel Survey," *Emerging Markets Finance and Trade,* pp. 1-18, 2026, doi: 10.1080/1540496X.2026.2630740.

[19] M. Rezagholizadeh, H. Jafari, and N. Kafi, "The Role of Technological Progress, Financial Market Risk, and Institutional Quality in the Impact of Natural Resource Revenues on Financial Development," *Iranian Economic Research,* vol. 30, no. 102, pp. 166-208, 2025.

[20] Y. Zhou and J. Chen, "Macro-prudential policy and bank systemic risk: Cross-country evidence based on emerging and advanced economies," *Emerging Markets Finance and Trade,* vol. 60, no. 5, pp. 1035-1047, 2024, doi: 10.1080/1540496X.2023.2266114.

[21] A. Zalbigi, S. Fattahi, and N. Ghaderi, "The Effect of Economic Policy Uncertainty on Stock Price Crash Risk in the Iranian Capital Market," *Journal of Financial Economics,* vol. 9, no. 3, pp. 33-50, 2023. [Online]. Available: https://www.jamv.ir/article_182869_77af8ca22289f61b363d8aa4b2f1370d.pdf.

[22] N. Salehi, R. Mohammadi, and M. Karimi, "Examining the Impact of Financial Constraints on Stock Price Crash Risk of Companies Listed in Tehran Stock Exchange," *Journal of Economics and Capital Market Management,* vol. 12, no. 4, pp. 87-108, 2023. [Online]. Available: https://civilica.com/doc/2083134/.

[23] G. M. Adhikari, N. Sapkota, D. Parajuli, and G. Bhattarai, "Impact of Green Banking Practices in Enhancing Customer Loyalty: Insights From Banking Sector Customers," *Financial Markets Institutions and Risks,* vol. 9, no. 1, pp. 195-215, 2025, doi: 10.61093/fmir.9(1).195-215.2025.