Section ECONOMETRICS

Gulara Rahimova
Baku State University
ORCID: 0000-0002-7820-5837
E-mail: gulararahim2021@mail.ru

Farhad Mirzayev
Baku State University
ORCID: 0000-0002-7820-5837
E-mail: farhad_1958@mail.ru

Ragsana Musayeva
Baku State University
ORCID: 0009-0002-2600-9786
E-mail: reqsanemusayeva8@gmail.com

https://doi.org/10.30546/200310.505.2026.1028

Econometrics of Risk: Theoretical and Methodological Foundations of Econometric Analysis of Risks

Abstract

Econometrics is the integration of economic theory, mathematical statistics, and empirical data, and in the insurance sector it is often regarded as the science of forecasting future outcomes based on historical information. In insurance, econometric methods are used to analyze and predict risks, losses, and financial indicators through mathematical and statistical models. This field lies at the intersection of econometrics and insurance practice, providing analytical tools for managing uncertainty and supporting evidence-based decision-making. Since insurance operations are inherently based on uncertainty, econometric analysis enables insurers to estimate future losses, calculate appropriate premium levels, and maintain financial solvency. Risk refers to the deviation of actual outcomes from expected results and, unlike uncertainty, can be measured and quantified. The insurance sector is a fundamental component of the modern digital economy, contributing to economic stability through risk distribution and financial protection mechanisms. This study examines the nature of risk factors affecting the insurance sector and explores their evaluation through econometric methods. The research analyzes the influence of macroeconomic and microeconomic indicators on insurance market performance using multivariate regression models. The findings indicate that inflation, gross domestic product, interest rates, and loss frequency significantly affect risk levels and the financial stability of insurance institutions.

Keywords: econometrics of risk, econometric model, correlation-regression analysis, risk forecasting, insurance sector.

Introduction

Econometrics of risk is a specialized branch of econometrics based on the quantitative modeling and statistical analysis of risk in various economic and financial contexts. This field incorporates multidisciplinary problems such as mathematical modeling, probability theory, and statistical inference for the evaluation of uncertainty, the measurement of risk exposure, and the forecasting of potential financial losses. This approach is widely applied in financial markets, insurance, macroeconomic policy, and corporate risk management. The main objective of econometrics of risk is the quantification of risk factors and the assessment of their impact on economic outcomes.

Econometrics of risk has emerged as a result of centuries of multidisciplinary research in mathematics, economics, and decision theory. According to Sakai's concept, its development is divided into six stages, and each stage is based on significant historical events that occurred.

1. Early Period (before 1700): This period is associated with the development of the foundations of probability theory. During this time, Blaise Pascal and Pierre de Fermat, in their 1654 correspondence concerning gambling games (the "problem of points"), formalized the concept of probability theory and laid the foundation for a new scientific discipline. In these studies, Pascal's work consisted of providing a scientific explanation for philosophical discussions related to early ideas of utility. Therefore, this research is known in science as Pascal's Wager Theory.

It was precisely during this period that mechanisms for risk distribution were established: Lloyd's Coffee House (1688), marine insurance, and stock exchanges such as the London Stock Exchange (whose predecessors existed from 1571, formally established in 1801) addressed practical risk problems related to trade and exploration, although the absence of formal economic foundations created limitations.

2. 1700–1880: This period is called the era of Bernoulli and Adam Smith. Daniel Bernoulli, in 1738, proposed the theory of expected utility to explain the St. Petersburg Paradox, replacing the expected monetary value with a logarithmic utility function. Adam Smith, in his 1776 work An Inquiry into the Nature and Causes of the Wealth of Nations, analyzed risky investments in markets and noted behavioral biases (tendencies). These cognitive and behavioral biases observed in decision-making under risk include individuals' inability to properly assess risk, making erroneous or irrational decisions under risky conditions, and systematic errors arising during information processing.

Major events of this period, such as the United States Declaration of Independence in 1776, the French Revolution in 1789, and the development of insurance institutions such as Tokio Marine Nichido (1879), increased the need for systematic risk assessment.

3. 1880–1940: This period is referred to as the era of Keynes and Knight. John Maynard Keynes, in his 1921 work A Treatise on Probability, distinguished between measurable risk and immeasurable uncertainty, and introduced the concept of "animal spirits." Frank Knight, in his 1921 book Risk, Uncertainty, and Profit, emphasized that profit arises from uninsured uncertainty. Additionally, World War I, the Great Depression, and the 1923 Great Kantō earthquake influenced the study of economic risks.

4. 1940–1970: This period is called the era of John von Neumann and Morgenstern. It is characterized as the period when game theory emerged. John von Neumann and Oskar Morgenstern, in their 1944 work Theory of Games and Economic Behavior, developed the expected utility approach. The second theory belonging to this period is Portfolio Theory. In this theory, Harry Markowitz introduced mean-variance analysis in 1952 to optimize the balance between risk and return (Markowitz, 1952). During the post-war development period, computational techniques and the Monte Carlo method accelerated risk modeling operations in the economy of the Cold War.

5. 1970–2000: This period in risk research is considered the era of Arrow, Akerlof, Spence, and Stiglitz. A characteristic feature of this era was information asymmetry, a situation in economic relations where one party possesses more and better-quality information than the other. According to this theory, in the insurance market, since the customer has more information about their health than the insurance company, high-risk individuals tend to purchase more insurance. In this context, George Akerlof's 1970 work The Market for Lemons introduced a new understanding of adverse selection and information asymmetry in markets (Akerlof, 1970). Consequently, volatility models were developed: Robert Engle created the ARCH model in 1982 (Engle, 1982), and Tim Bollerslev developed the GARCH model in 1986, enabling the measurement of time-varying volatility (Bollerslev, 1986). The 1973 oil crisis, the Chernobyl disaster in 1986, and the Dissolution of the Soviet Union exposed weaknesses in these models.

6. From 2000 to the Present Day: This period is considered the modern era in the development of risk studies. During this time, systemic risks emerged. For example, the 2008 financial crisis revealed the limitations of traditional risk measurement approaches and increased attention to systemic risk within financial and insurance markets (Cummins & Weiss, 2014). New instruments for econometric analysis of risks in insurance were developed: machine learning, extreme value theory, and Bayesian networks began to be widely used in modeling extreme risks. During this period, regulatory mechanisms were also developed, with special attention given to stress testing and liquidity risk in Basel III and Basel IV standards.

Methodological Foundations of Econometric Risk Analysis

Econometric risk analysis is one of the important scientific approaches for assessing risks in the insurance sector. Through econometric models, this approach enables the identification of functional relationships existing between risk factors and their outcomes. The application of econometric techniques contributes significantly to the efficiency and performance of insurance firms through improved risk management and financial intermediation activities (Cummins et al., 2009). The main objective of econometric risk analysis is to move away from subjective and intuitive assessments and ensure that risks are expressed through more precise quantitative indicators. This makes it possible to account for the multidimensional and complex nature of risks. Modern financial risk assessment methods are also closely linked to portfolio diversification principles (Markowitz, 1952) and asset valuation techniques developed in financial economics (Black & Scholes, 1973).

Furthermore, studies indicate that insurance sector development contributes positively to economic growth and financial stability (Outreville, 2013). Research on European insurance markets has also demonstrated that regulatory conditions, market competition, and business environments significantly influence the productivity and efficiency of insurance companies (Eling & Schaper, 2017). The following models are used in econometric risk analysis:

Linear Regression Models

These models allow the studied relationship to be expressed in the form of an analytical equation. In such models, dependent and independent (explanatory) variables are determined in advance, and a linear relationship is established between the risk indicator and the factors affecting it. This is expressed in the following form:

$$Y=\beta_0+\beta_1X_1+\cdots+\beta_nX_n+\varepsilon.$$

In the insurance context, the fundamental econometric equation describing insurance loss (Y) through the simplest linear model is represented as:

$$Y_i=\beta_0+\beta_1X_{1i}+\beta_2X_{2i}\cdots+\beta_kX_{ki}+\varepsilon_i.$$

Where:

- Y_i – Expected amount of loss;
- X – Risk factors (for example, age, engine capacity);
- B_i – The impact strength of each factor;
- ε_i – Error term. In insurance, this term is often very large because a significant part of risk is inherently random. The econometrician's task is to identify regularities hidden within this error component.

Generalized Linear Models (GLM)

GLM is considered the most powerful econometric tool in calculating insurance premiums because insurance losses do not follow a normal distribution.

GLM consists of three main components:

1. Random Component: The distribution of the dependent variable (Y), such as Poisson, Gamma, or Pareto.
2. Systematic Component: Formed from the linear combination of explanatory variables (Xi).
3. Link Function: A function that mathematically connects the expected value with the linear predictor (for example, log, logit).

For example, when applying a log-link function for automobile insurance, the model becomes:

$$ln(E[Y]) = \beta_0 + \beta_1 Yaş + \beta_2 Region,$$

Here, $\beta 0, \beta 1, \beta 2$ represent the impact strength of the factors.

Logistic Regression Models

These models account for the complexity and multidimensionality of risks and are particularly used in evaluating the probability of an event occurring. For example, this model is applied to determine the probability of an insurance event taking place. This method enables the transition from experts' subjective assessments to formal probability indicators and helps forecast the probability of project failure.

Integration of Econometric Risk Analysis and Time Series Models

Time series analysis and risk econometrics are among the main analytical approaches used in modern econometrics and financial economics for measuring, modeling, and forecasting risks. This field particularly serves to establish scientifically grounded decision-making processes under uncertainty.

A time series is a set of observations recorded sequentially over specific time intervals.

Economic and financial indicators (prices, revenues, interest rates) are usually presented in time-series form. Main components of time series include:

- *Trend* – Long-term tendency of growth or decline;
- *Seasonal component* – Periodically repeating changes;
- *Cyclical component* – Fluctuations related to economic cycles;
- *Random component* – Unexplained stochastic variability.

In risk econometrics, time series are used for:

- Analyzing changes in risks over time;
- Forecasting future risk levels;
- Evaluating financial stability;
- Studying the dynamics of insurance payments.

For example, insurance companies can analyze annual changes in claim payments to forecast future risks.

The following classical and modern time series models are widely used for risk analysis:

AR (Autoregressive) Model

$$Y_t=\alpha+\beta Y_{t-1}+\varepsilon_t$$

In this model, the current value depends on the previous period's value.

Where:

- Y_t – Dependent variable showing the outcome at time *t* during analysis. In insurance examples, this may represent the number of insurance claims, the volume of insurance premiums, or the amount of loss as a risk indicator;

- α *(intercept)* – Constant term representing the expected value of Y_t when $Y_t -1 = 0$. In insurance, α reflects the model's baseline level, meaning the expected minimum risk/premium value without considering historical observations. The expression risk/premium is mainly used in insurance and financial contexts: *Risk* – The probability of a certain event occurring or the potential loss caused by that event. For example, the risk of being involved in a car accident in automobile insurance, or illness/death risk in life insurance. *Premium* – The amount paid by the insured party to the insurance company under the insurance contract. This amount is determined according to the level of risk, coverage scope, and insurance duration. Thus, in insurance theory, the concept of risk/premium usually refers to "a premium corresponding to the insurance risk": the higher the risk, the higher the premium is likely to be.

- β – Autocorrelation/effect parameter measuring the influence of changes in $Y_t -1$ on Y_t.

- If $\beta > 0$, a high risk/premium in the previous period indicates a higher probability of remaining high in the current period.

- If $\beta < 0$, high risk in the previous period may lead to a lower level in the current period. In insurance, β is important for measuring the persistence or trend of risks.

- ε_t – Residual (error) term representing random changes that the model cannot explain. In insurance, this includes impacts arising from unexpected events such as accidents, natural disasters, and similar occurrences.

This is an *AR(1) model***,** in other words, a first-order autoregressive model. In more complex insurance analysis, **εt** may exhibit a heavy-tailed distribution because rare but severe risks exist. The parameters of these risks are estimated using OLS (Ordinary Least Squares) or MLE (Maximum Likelihood Estimation) methods.

OLS (Ordinary Least Squares) is the most commonly used method in statistics and econometric modeling for estimating parameters related to the error term. Its objective is to minimize the sum of squared differences between observed quantities and the quantities predicted by the model. Suppose we have a simple linear regression model:

$$Y_i=\alpha+\beta X_i+\varepsilon_i\,.$$

Where: Y_i – Dependent variable; X_i – Independent variable; ε_i – Errors; α, β – Parameters we want to estimate.

In the OLS method, we need to minimize the following objective function for α and β:

$$\min\sum_{i=1}^{n}\left(Y_i-\alpha-\beta X_i\right)^2 .$$

Characteristics of OLS: Simple and fast to apply; Under the assumption of normally distributed errors, it is considered the BLUE (Best Linear Unbiased Estimator)**.**

MLE (Maximum Likelihood Estimation) is a parameter estimation method based on probability theory. Its purpose is to identify the parameter values that maximize the likelihood of observing the given data. Suppose our observations Y_i follow a certain probability distribution:

$$\theta_{MLE}=\arg\max_{\theta}\prod_{i=1}^{n}f(Y_i\,|\,\theta)\,.$$

Where:

-$(X_1, \ldots, X_n)$ are observations drawn from the same distribution. This distribution is characterized by parameter θ***;***

- $f(Y_i \mid \theta)$ is the density function, or in the discrete case, the probability function;

-θ is the parameter we aim to estimate. That is, we seek the value of θ that maximizes the probability of the observed data.

Characteristics: It is highly flexible and can be applied to any distribution; for large samples, the MLE estimator is consistent (as the number of observations increases, its value approaches the true parameter, and the estimation error decreases), asymptotically normal, and efficient. The OLS method, when errors are normally distributed, can be considered a special case of MLE. A simple comparison table and graph are shown below (Table 1, Figure 1):

Table 1.

Characteristic	**OLS**	**MLE**
Objective	Minimize the sum of squared errors	Maximize the likelihood (probability)
Main Assumption	Errors have zero mean and are homoscedastic	The distribution form must be known
Advantage	Simple and fast	More flexible, suitable for various distributions

MA (Moving Average) Model (Moving Average Calculation Method). This model is a statistical method used in time series analysis to smooth short-term fluctuations in indicators and to observe the overall trend more clearly. In other words, MA means that the average of a certain number of consecutive observations is calculated, and in the next period this window “moves forward.” A simple moving average is calculated as:

$$MA_t = \frac{Y_t + Y_{t-1} + \ldots + Y_{t-n-1}}{n}$$

In more complex situations, the MA(q) model is also written as:

$$Y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots + \theta_q\varepsilon_{t-q}$$

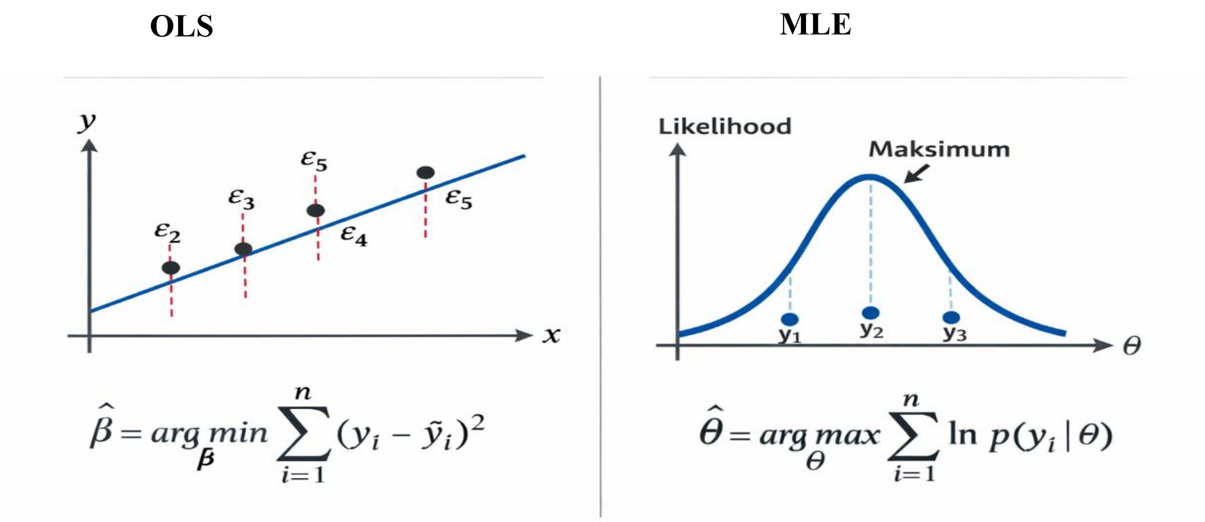

Figure 1. Graphical representation of OLS versus MLE.

Where:

- Y_t – Time series variable (parameter) at time t;
- μ – Mean or constant value;
- ε_t – White noise or random errors;
- $\theta_1, \theta_2, \ldots, \theta q$ – Parameters of the MA model;
- q – Order of the MA model (how many lagged error effects are considered).

In the simple case, the MA(1) model is expressed as:

$$Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}.$$

Here, Y_t is the value of the current period, and n is the number of periods over which the average is taken. Previous errors are considered in this process. For example, let us calculate a 3-day moving average (Table 2): Example data:

Table 2

1	10
2	12
3	14

For Day 3: MA=(10+12+14)/3=12, then (12 + 14 + new value) / 3.

In applications within the insurance sector, MA reduces random fluctuations, helps identify trends, and is used in forecasting risks.

ARMA Model

This is a combination of AR and MA models and is used for evaluating stationary time series. Stationarity of time series is an important condition in econometric analysis. Non-stationary series are transformed into another form through certain mathematical operations, such as differencing, in order to build the model more accurately and improve data properties.

ARIMA Model (Autoregressive Integrated Moving Average). The ARIMA model is widely used in the insurance sector to analyze time series of risks and revenues. This model is applied to non-stationary series and is generally represented as:

ARIMA(p,d,q),

Where:

p – Order of autoregression; d – Level of differencing; q – Order of moving average.

This model combines autoregressive (AR), integrated (I), and moving average (MA) components to explain time series dynamics. It is effective in short-term forecasting.

Structure of the ARIMA Model. The components of the ARIMA(p,d,q) model can be expressed in formula form as follows:

1. AR(p) – Autoregressive Component: Shows the effect of past observations on the current value.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t .$$

2. I(d) – Integration Component: If a time series is non-stationary, a differencing operation ($\Delta^d Y_t$) is applied to make it stationary.

3. MA(q) – Moving Average Component: Shows the effect of past random errors (residuals) on the current value.

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} .$$

When all components are combined, the ARIMA(p,d,q) model is formed.

Application of the ARIMA Model in the Insurance Sector

1. Claims Forecasting: Analyzes the time series of monthly or annual claims for a specific type of insurance. For example, in automobile insurance claims: The AR component reflects the impact of past claims; The MA component reflects unexpected events.

2. Risk Management: The ARIMA model helps insurance companies forecast the magnitude of future risks and allocate capital for those risks. It is particularly useful in modeling heavy-tailed and non-stationary risks.

3. Premium Setting: Forecasting future payments based on past claims is used in determining premiums. For example, annual insurance premiums can be optimized using the ARIMA model.

4. Comparison of Different Insurance Portfolios: By modeling time series for different types of insurance, it is possible to analyze the distribution of risks across portfolios.

Simple Practical Example (Zolkin A. L., 2025)

$$\Delta Y_t = \phi_1 \Delta Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} .$$

Where

- $\Delta Y_t = Y_t - Y_{t-1}$, ϕ_1 – Effect of past monthly changes;
- θ_1 – Past effect of random changes.

Based on this model, a company can forecast the number of insurance claims for the next month.

Now let us write Python program code for an ARIMA model built on simple insurance claims time series data and construct the forecast graph.

```
# Required libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA

# 1. Sample insurance claims data (monthly)
np.random.seed(42)
months = pd.date_range(start='2022-01', periods=36, freq='M')

# Base trend + random noise
claims = 50 + np.cumsum(np.random.randint(-5, 6, size=36))
data = pd.Series(claims, index=months)

# 2. Plot of the time series
plt.figure(figsize=(10,4))
plt.plot(data, marker='o')
plt.title('Monthly Insurance Claims')
plt.xlabel('Month')
plt.ylabel('Number of Claims')
plt.grid(True)
plt.show()

# 3. Building the ARIMA model (p=1, d=1, q=1)
model = ARIMA(data, order=(1,1,1))
model_fit = model.fit()

# 4. Forecast (next 12 months)
forecast = model_fit.forecast(steps=12)
forecast_index = pd.date_range(
    start=data.index[-1] + pd.offsets.MonthEnd(1),
    periods=12,
    freq='M'
)
forecast_series = pd.Series(forecast, index=forecast_index)

# 5. Main graph + forecast
plt.figure(figsize=(10,4))
plt.plot(data, label='Observed Claims', marker='o')
plt.plot(
    forecast_series,
    label='ARIMA Forecast',
    marker='x',
    color='red'
)
plt.title('ARIMA(1,1,1) Forecast of Insurance Claims')
plt.xlabel('Month')
plt.ylabel('Number of Claims')
plt.legend()
plt.grid(True)
```

```
plt.show()
```

Explanation of the code:

1. `data` – In our example, this represents a 3-year (36-month) sample of insurance claims.
2. `ARIMA(data, order=(1,1,1))` – ARIMA(1,1,1) model:
   - p = 1 → AR(1);
   - d = 1 → first differencing;
   - q = 1 → MA(1).
3. `forecast(steps=12)` – Produces a forecast for the next 12 months (Figure 2).
4. `plt.show()` – Used to display the graph on the screen.

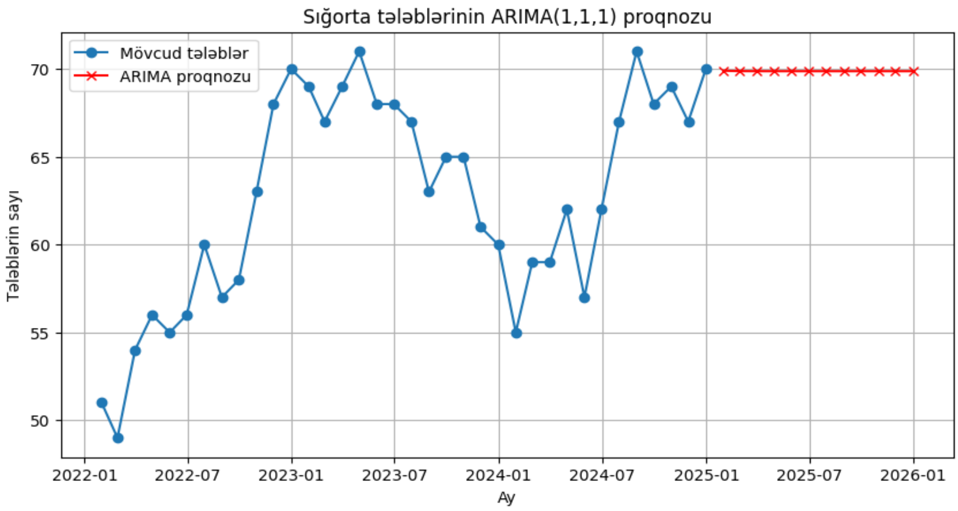

Figure 2. Forecast of insurance claims based on the ARIMA (1,1,1) model built in Python.

In the graph, the blue line represents the current monthly claims, while the red line shows the ARIMA forecast for the next 12 months. The graph demonstrates that the model follows the existing trend and predicts increases and decreases in claims for the future.

For real insurance data and heavy-tailed risks, a more advanced ARIMAX/non-normal model can also be built, which would provide a more realistic forecast (Figure 3).

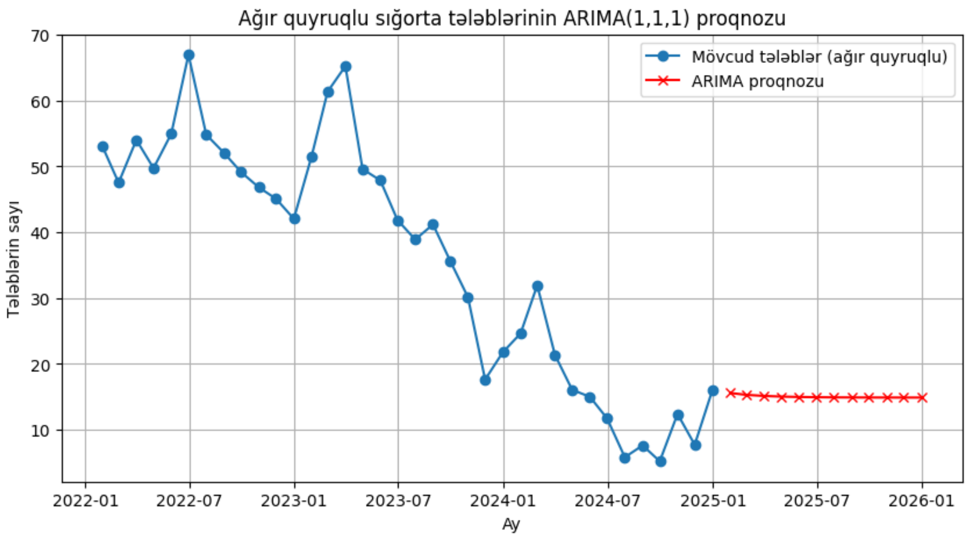

Figure 3. A more realistic scenario: ARIMA(1,1,1) forecast for heavy-tailed insurance claims.

The blue line shows the actual monthly claims with a heavy-tailed distribution, and the red line shows the ARIMA forecast for the next 12 months. The graph clearly shows that there are large jumps (heavy tails) in the claims, and the model takes them into account in the forecast.

Volatility models (fluctuation models, such as GARCH) and multivariate volatility models

1. These models are used to measure the time-varying nature of financial risks and are widely applied in assessing the risk level of insurance portfolios. In stochastic volatility (SV) models, volatility is treated as a separate latent stochastic process. These models play a crucial role in forecasting time-varying risks, measuring financial risks, and pricing derivative instruments (Babicheva I.V., “Lan,” 2024).

In econometric analysis of insurance risks, volatility models are highly effective because they allow measurement and forecasting of risk variability (variance) over time. Volatility is the change in variance over time:

$$Var(Y_t) = \sigma_t^2 .$$

If σ_t^2 is non-stationary, heteroskedasticity arises, and classical models such as OLS are insufficient for risk analysis.

As a primary example of volatility models, the ARCH model developed by Engle in 1982 can be cited. Mathematically, the model can be expressed as follows:

$$\varepsilon_t \sim N(0, \sigma_t^2),$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 .$$

This formula shows that current risk is directly proportional to past squared errors, meaning a large shock will cause higher future risk.

Volatility models in insurance econometrics check risk levels, measure variability, and are especially important for heavy-tailed and unstable risks. They are most effective when used together with ARIMA.

In 1986, Bollerslev developed the GARCH model, which is more realistic and “memory-aware.” If the volatility depends on its past values, the model is called a “memory” volatility model.

In other words, if there is a large fluctuation (shock) in the market or insurance claims, its effect does not disappear immediately but persists for a period. If risk is high today, it is likely to remain high tomorrow, and if low today, it tends to stay low tomorrow. This phenomenon is called volatility clustering.

ARCH(1,1) model. In risk econometrics, this model is particularly important:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 .$$

Here, σ_t^2 is the current volatility, σ_{t-1}^2 is the effect of past shocks, and ε_{t-1}^2 is the coefficient of past volatility. The larger β_1 is, the longer the memory of volatility—i.e., past risk influences future risk for a longer period. The GARCH model shows risk peaks.

This model captures the time-varying nature of volatility and is particularly used in assessing insurance and financial risks. ARCH models account for conditional variance based on past shocks and help detect volatility clustering. GARCH extends this approach by including lagged variances. Models like EGARCH consider asymmetry (e.g., leverage effect).

In general, when natural disasters—such as strong earthquakes or floods—occur, the risks they create do not disappear immediately, and risk "retains memory." That is, during a severe flood or inundation, consecutive accidents or aftershocks can happen over several days, and the likelihood of such events occurring in subsequent periods remains high. For example, on October 22 and 29, 2024 (5,021 insurance claims) and March 27 and 28, 2026 (4,586 insurance claims), heavy rains in Baku and other regions of Azerbaijan caused floods and damage that resulted in substantial consecutive losses for people. According to statistical data, during these rainy periods, the number of car accidents increased 3.4 times. To address these issues, the insurance sector had to allocate substantial funds. This created prolonged instability characteristic of a crisis period in the financial and insurance markets. Such situations can be represented using ARIMA models for mean memory and GARCH models for volatility memory. The econometric formula is as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 .$$

Here, σ_{t-1}^2 is a quantity that characterizes past risk, essentially representing "memory."

If a model has volatility memory, then risks cluster—that is, they occur in groups; classical models with constant variance give misleading results; more capital reserves are required, and in such cases, models like GARCH should be used.

Thus, we conclude that *"volatility has memory"* means that the variability of risk is not random: past risks inevitably influence future risks, and this effect gradually diminishes over time. In other words, if risk has memory:

1. Risks will cluster—consecutive periods will be either high or low;
2. Prediction becomes possible; and
3. The insurance company will be able to calculate reserves and premiums more accurately.

Another example of such a model is the EGARCH model.

$$\ln\left(\sigma_t^2\right) = \omega + \beta \ln\left(\sigma_{t-1}^2\right) + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| .$$

The main feature is that it accounts for asymmetric effects, meaning positive and negative shocks have different impacts. This additional parameter creates a more realistic model (Sharapov Yu.V., Stovba E.V., Sharapova N.V., 2025).

TGARCH / GJR-GARCH model

This model treats bad and good shocks differently (bad news creates higher risk, while good news has a weaker effect).

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma D_{t-1} \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 .$$

Its application in insurance risk: when calculating reserves for future claims, if volatility is high, more capital is required.

The application of this model in insurance risk lies in the fact that, when calculating insurance reserves, the company sets aside funds for future claims, and if volatility is high, it is taken into account that more capital is required.

Risk measures, such as Value at Risk (VaR) and Expected Shortfall (ES), are calculated using volatility models and are especially important for heavy-tailed distributions.

If risk is highly variable, reinsurance is used. Reinsurance is a process in which an insurance company transfers a portion of the risk it has assumed to another insurance company.

The calculation of insurance premiums depends on the insurance system (mandatory state social insurance, voluntary insurance, etc.) and the applicable legislation. The general principle is the same: the insured income or risk is multiplied by a certain rate. Insurance premiums are usually calculated as:

Insurance premium = Insurance base × Rate (percentage)

For mandatory state social insurance in Azerbaijan:

a) For the employee – a certain percentage is deducted from the employee's salary, e.g., 3% (or it may vary according to legislation);

b) For the employer – the employer pays an additional percentage, e.g., 22% (or differential rates may apply).

Example: If the salary is 1500 AZN:

- Employee: 1500 × 3% = 45 AZN
- Employer: 1500 × 22% = 330 AZN
- Total amount: 375 AZN

For voluntary insurance (e.g., life insurance, car insurance), the calculation is more complex and depends on factors such as risk level (age, health, driving experience, etc.), insured amount, duration, statistics, and probability calculations.

A simple approach is:

Premium = Probability of risk × Insurance amount + additional costs.

Additional factors include minimum and maximum insurance base, discounts or additional charges, and sector-specific rates (construction, public sector, etc.).

Use of ARIMA + GARCH models together.

The combined use of ARIMA and GARCH models is widely applied to model both the mean dynamics and volatility of time series. This approach is particularly effective for financial data (e.g., income formation). The ARIMA model captures the mean level of the time series, while the GARCH model captures the variance (volatility) of the residuals. In other words, the forecast is first made with ARIMA, and then GARCH is applied to the ARIMA residuals.

For the ARIMA model (mean equation), we can use the following form:

$$y_t = \mu + \sum \phi_i y_{t-1} + \sum \theta_j \varepsilon_{t-j} + \varepsilon_t \,,$$

Here, ε_t are the residuals.

For the GARCH model (variance equation):

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 .$$

Here, σ_t^2 is the conditional variance, and

$$\varepsilon_t \sim N(0, \sigma_t^2)$$

are the residuals.

Step-by-step implementation consists of the following sequential stages:

1. Checking for stationarity;
2. Selecting the ARIMA model;
3. Examining the residuals;
4. Building the GARCH model;
5. Making the forecast.

For example, in the case of financial returns: the ARIMA model is applied to estimate the expected value of returns, the GARCH model is used to assess risk, and volatility shows the variability over time. As a result, a model is constructed that determines both the return forecast and the risk measure.

Advantages: It is important in financial risk analysis because it allows constructing a model more consistent with real data. Disadvantages: Model selection is complex, parameters are sensitive, and computation can be intensive.

VAR (Value-at-Risk) models

This model measures the maximum possible loss with a certain probability, allows the simultaneous analysis of interactions among several variables, and is important for assessing systemic risks. The Expected Shortfall (ES) model takes a more conservative approach than VaR and accounts for extreme losses.

Risk indicators (VaR, expected loss) and quantile methods are widely used.

Econometricians evaluate risk measures in insurance value at risk (VaR) and expected shortfall (ES) using both parametric and non-parametric methods.

Extreme value theory is used to model “tail” risks and allows the assessment of potential losses at high quantiles.

Quantile regression, in turn, directly models the conditional quantiles of returns, enabling the calculation of the maximum expected loss at a specified confidence level.

Conclusion

The theoretical and methodological foundations of econometric analysis of risks provide scientifically grounded tools for better understanding, measuring, and managing risks. Using econometric models, the level of risk—for example, the probability of losses is quantified, variability (volatility) and uncertainty are expressed numerically, meaning that risk is no longer “approximate” but becomes a measurable indicator. Econometric analysis with models such as ARIMA, GARCH, and others allows the forecasting of future losses, claim counts, and financial damages, thereby strengthening decision-making.

Econometric models applied to identify risk factors reveal which factors increase risk during economic crises, interest rate calculations, and market volatility, as well as which variables have stronger impacts. Based on the results obtained, risks are managed: insurance tariffs are set correctly, reserve funds are calculated, and risk mitigation strategies are developed. Companies can make more informed decisions, minimize losses, increase profitability, and thus optimize the decision-making process. In conducting econometric risk analysis, we are able to measure, explain, forecast, and manage risks.

References

1. *“Econometrics: A Course of Lectures in Diagrams, Tables, and Examples”, author: I. V. Babicheva, publishing house “Lan”, 2024.*
2. *“Economic-Mathematical Methods: Practical Guide”, authors: Yu. V. Sharapov, E. V. Stovba, N. V. Sharapova, V. M. Sharapova, publishing house “Dashkov & Co.”, 2025.*

3. *"Fundamentals of Building Economic Models in Excel", authors: Yu. E. Voskoboynikov, I. N. Mukhina, publishing house "Lan", 2026.*
4. *"Programming of Econometric Models", author: A. L. Zolkin, 2025.*
5. *Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism.* The Quarterly Journal of Economics, 84*(3), 488–500.* *https://doi.org/10.2307/1879431*
6. *Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities.* Journal of Political Economy, 81*(3), 637–654.* *https://doi.org/10.1086/260062*
7. *Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity.* Journal of Econometrics, 31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1*
8. *Cummins, J. D., & Weiss, M. A. (2014). Systemic risk and the U.S. insurance sector.* The Journal of Risk and Insurance, 81*(3), 489–528.* *https://doi.org/10.1111/jori.12039*
9. *Cummins, J. D., Dionne, G., Gagné, R., & Nouira, A. (2009). Efficiency of insurance firms with endogenous risk management and financial intermediation activities.* Journal of Banking & Finance, 33*(1), 180–192.* *https://doi.org/10.1016/j.jbankfin.2008.07.015*
10. *Eling, M., & Schaper, P. (2017). Under pressure: How the business environment affects productivity and efficiency of European life insurance companies.* European Journal of Operational Research, 258*(3), 1082–1094.* *https://doi.org/10.1016/j.ejor.2016.08.070*
11. *Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation.* Econometrica, 50*(4), 987–1007.* *https://doi.org/10.2307/1912773*
12. *Markowitz, H. (1952). Portfolio selection.* The Journal of Finance, 7*(1), 77–91.* *https://doi.org/10.1111/j.1540-6261.1952.tb01525.x*
13. *Outreville, J. F. (2013). The relationship between insurance and economic development: 85 empirical papers for a review of the literature.* Risk Management and Insurance Review, 16*(1), 71–122.* *https://doi.org/10.1111/rmir.12004*