

International Journal of Advanced Statistics and Probability

Website: www.sciencepubco.com/index.php/IJASP

Research paper



# The main determinants of the absorptive capacity of the Egyptian insurance market using statistical analysis

Dr. Nermin Saad Mohmed Fahmi \*

Department of Statistics Faculty of Science King Abdu-Aziz University \*Corresponding author E-mail: dr.nerminfahmy@yahoo.com

## Abstract

This research aims to use a statistical and economic model that takes into account all the emerging economic, social and financial variables in the framework of the new world order and the new general situation to determine the most important basic determinants affecting the volume of demand and the supply of insurance policies for each branch of insurance separately. To maximize its financial, economic and social return, that is, to determine the factors affecting the absorptive capacity of the Egyptian insurance market using some statistical models. It was found that the best regression models that can be used to describe the relationship between the number of insurance policies and the economic variables used in the study is the semi-logarithmic model, and the size of insurance amounts is related to a group of economic variables at the national level, including: population - size of the labor force - birth rates, and it is also related to per capita GDP.

Keywords: Linear equation in matrix; Half Logarithmic Equation; Logarithmic equation; The demand function; Capita gross domestic product.

## 1. Introduction

The insurance industry plays an effective role in contemporary economies, and it is clear that the insurance industry in industrialized countries has participated in economic growth sectorally and geographically, as it has intertwined with the sectors of industries, petroleum, mining, agriculture, transportation of all kinds, internal and external trade, human activities for all age groups and all risks. Insurance, as an integrated system for risk management and distribution, was able to affect savings and investment, given that insurance and reinsurance companies are among the most important financial institutions in the state's economy.

The insurance market in Egypt used to be operating under certain environmental, political, social and economic conditions, but now the world and its economic, political, social and cultural system are in a continuous development and in a transitional phase with repercussions and repercussions on all economic and social variables and all economic and social sectors. This study is concerned with estimating the effective demand for various types of insurance services in the Egyptian market, whether in the field of personal or property and liability insurance, and determining the most important factors affecting the volume of demand for various types of insurance, which will enable us to develop the supply side, whether in terms of promotional policies adopted or developing The quality of the insurance product, so that the insurance market can be divided into homogeneous consumer groups according to the characteristics of the insured.

This research aims to use a statistical and economic model that takes into account all the emerging economic, social and financial variables in the framework of the new world order and the new general situation to determine the most important basic determinants affecting the volume of demand and the supply of insurance policies for each branch of insurance separately. To maximize its financial, economic and social return, that is, to determine the factors affecting the absorptive capacity of the Egyptian insurance market using some statistical models.

# 2. The proposed model

The method of multiple regression will be used to study the relationship between a dependent variable and a group of independent variables and arrive at an equation that can be used in predicting the values of the dependent variable in terms of the values of the (independent) explanatory variables and their coefficients in the regression equation.

The method of multiple linear regression will be used to arrive at a linear relationship between the number of new insurance policies Y1 and the values of insurance amounts Y2 as two dependent variables and a set of independent variables including the following:

- Time (x1).
- Total population (x2).
- Birth rate (in thousand) (X3).
- Marriage contract rate (in thousand) (X4).
- GDP per capita (X5)
- Purchasing power of money (based on the year 95/96) (X6).



Copyright © Dr. Nermin Saad Fahmi. This is an open access article distributed under the <u>Creative Commons Attribution License</u>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

- Total work force (x7).
- Per capita share of social insurance contributions (X8).
- Savings volume (x9).
- Number of private insurance funds (X10).
- Total Insurance Premiums (X11).
- Number of producers (X12).

#### 2.1. Mathematical formulation of the model

The mathematical formulation of the model means the number of equations contained in the model, the degree of linearity of the model, and the degree of homogeneity of the equations.

In its estimation of the demand functions for life insurance, the paper will use three standard equations, which are the linear equation using absolute values, the half-logarithmic equation and the logarithmic equation:

• Linear equation in matrix form:



• Half Logarithmic Equation in Matrix Form:

$\begin{bmatrix} b_{0} \\ b_{1} \end{bmatrix}$		$\begin{bmatrix} n \\ \sum X \end{bmatrix}$	$\sum X_{1}$ $\sum X_{1}^{2}$	  •	• •	  •	$\begin{bmatrix} \sum X_{13} \\ \sum X_{1}X_{13} \end{bmatrix}$	$\begin{bmatrix} \sum \ln y \\ \sum X_{\perp} \ln y \end{bmatrix}$
	_	•						
•	_	•						
		$\sum X_{14}$	$\sum X_{1}X_{2}$				$\sum X_{13}^2$	$\left\lfloor \sum_{x} X_{13} \ln y \right\rfloor$

• Logarithmic equation in matrix form:

$\begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$		$\int_{\sum \ln X_1}^{n}$	$\frac{\sum \ln X}{\sum (\ln X)_{1}^{2}}$	•	•	•	•	•	•	•	$\frac{\sum \ln X_{14}}{\sum \ln X_{1}X_{14}} \bigg ^{-1}$	$\left[\begin{array}{c} \sum \ln y \\ \sum \ln x_1 y \end{array}\right]$
				·	•	•		·		·		
	=											
											•	
.		•									•	
.		-									•	
		$\sum_{13}^{13} \ln X_{13}$	$\sum \ln X_{1}X_{2}$								$\sum (\ln X)_{13}^2$	$\left[\sum \ln x_{13} \ln y\right]$

Where the dependent variables are Y1, the number of insurance policies, and Y2, the values of insurance amounts, the researcher will differentiate between the three attempts according to the value of the coefficient of determination in each case in addition to the substance of the estimated relationship as a whole and the substance of the estimated transactions in addition to the suitability of the transactions to the requirements of economic theory.

The formation of the proposed statistical model requires converting the mathematical relationship into a statistical or probability relationship, by adding the error term element to the mathematical model. The random error component expresses that irregular part in the behavior of the variables to be predicted, which results from several sources:

- The explanatory variables that are removed from the model during the estimation process.
- Measurement, viewing and assembly errors.
- Model drafting errors.

The following is a statement of the statistical functions of the demand for life insurance documents that will be assessed:

a) The demand function on the number of documents

 $y_1 = b_0 + b_1 x_1 + \dots + b_{13} x_{13} + e_1$ 

b) The demand function for the amount of documents

 $y_{2} = b_{0} + b_{1}x_{1} + \dots + b_{13}x_{13} + e_{2}$ 

Where e<sub>1</sub>, e<sub>2</sub> are random variables that represent the random error component in the estimation of the two functions of ordering demand, and the researcher will apply the Least Squares Method, which is based on the following assumptions:

- 1) The error component is a random variable that takes probability values.
- 2) The expected value of the random variable = zero That is, [E (e) = 0]
- 3) A variance of  $e_i$  around its arithmetic mean is constant for all values of the variable explained, i.e.  $\left[E(e_i^2) = \sigma_e^2\right]$
- 4) The values of the random variable  $e_i$  is independent of each other, i.e.,  $cov (e_i e_j) = 0$
- 5) The distribution of the random variable  $e_i$  is normal, its arithmetic mean = zero and its variance  $\sigma_{\epsilon}^2$
- 6) The random variable  $e_i$  is independent of the explained variables in the model.

## 2.2. Model estimation

Estimating the model requires collecting data on all the explained and dependent variables and preparing them in order to become suitable for applying the statistical method on them to reach the required results. Therefore, the researcher is exposed to the following:

• How to prepare the data

The statistical method that will be applied to the demand functions.

• Preparing statistical data

It is known that the accuracy of the results obtained depends on the degree of accuracy in gathering data from its various sources, and the degree of representation of that data for the different variables included in the study. The lack of data is one of the biggest problems when applying the model as it is easy to create models of all types and sizes that can be dealt with in different ways, but finding different data suitable for a particular model is another issue because in general we find that the data is not available or published in the required form, which leads to the work of some Modifications or use of modified data.

• The statistical method used to estimate the model parameters

Methods for estimating the model parameters of economic relations are divided into two groups:

• Methods for the one equation method

The most important of these are the Regular Least Squares Method (OLS), the Indirect Least Squares Method (ILS), the Two-Stage Least Squares Method (2SLS), and the Maximum Likelihood Limited Information Method. These methods are used if prediction is the main goal of the model.

• Methods for the following equations

The most important of which is the three-stage least squares method and the Full Information Maximum Likelihood method, and since the main objective of the model is to estimate the demand functions for use in forecasting, the single equation method will be used to prepare a statistical model that contributes to estimating the demand for individual life insurance documents in the Egyptian market.

In the statistical analysis, the researcher will rely on the experimental method, and through it, one can start with a simple model depending on the causal considerations, then make gradual adjustments based on the resulting statistical indicators. By adding new variables to the relationship and studying its statistical essence every time.

- Depending on the experimental method, the following three attempts will be made
- a) Matching a linear regression relationship using the absolute values of the observations
- b) Matching a regression relationship using the logarithmic formula
- c) Matching a regression relationship using the semi-logarithmic formula

These attempts provide better and greater opportunities to choose the best model based on statistical and economic criteria to reach the most accurate results by applying the methods of gradual multiple regression and elimination from the back.

• Stepwise multiple regression method

This method aims to obtain a regression equation that does not necessarily contain all the explained variables but rather expresses the problem under study in the best possible expression and the gradient regression method is performed in several steps, through which an intermediate regression equation is obtained by adding or excluding one of the explained variables according to the degree of its addition in reducing unexplained variance so that the addition or exclusion of the variable contributes to a greater improvement in the accuracy of reconciliation and the substance of the relationship.

The correlation matrix is of great importance in this method because it contains partial correlation coefficients between the dependent variable and each independent variable, which shows the degree of correlation between the dependent variable and each independent variable separately, and therefore the variables that can enter or exit the relationship are selected according to the degree of their correlation. With the dependent variable.

• Backward elimination method

This method is performed on several steps. In each step an intermediate regression equation is obtained by deleting one variable so that the elimination of this variable leads to a greater improvement in the accuracy of the reconciliation.

While the estimated coefficients represent the best values for reconciling the equation with the remaining variables in the model.

## **3. Practical application**

The results will be presented here with regard to estimating the volume of demand for various types of insurance in the Egyptian market. A set of mathematical models were used in which the dependent variable is represented in the volume of demand with different independent variables in each case. The size of the order was represented as a dependent variable by:

- The number of insurance policies (the branch of life): whether by using the number of documents issued in the insurance market or by using insurance funds.
- Direct subscription premiums (General Insurance Branch).
- Subscriptions to private insurance funds.
- All kinds of regression method will be used to do these statistical analyzes to estimate the actual demand for insurance in the Egyptian market.

As previously mentioned, the multiple regression method is used to study the relationship between a dependent variable and a group of independent variables and arrive at an equation that can be used in predicting the values of the dependent variable in terms of the values of the (independent) explanatory variables and their coefficients in the regression equation. The regression method and Curve Fitting methods were used to estimate the demand for various types of insurance through the dependent variables that we have referred to.

a) Insurances of persons

In this study, multiple linear regression method was used to arrive at a linear relationship between the number of new insurance policies (Y1) and the values of insurance amounts (Y2) as two dependent variables and between a set of independent variables that include the following:

- Time (x1)
- Total population (x2)
- Birth rate (in thousand) (X3)
- Marriage contract rate (in thousand) (X4)
- GDP per capita (X5)
- Purchasing power of money (based on the year 95/96) (X6)
- Total work force (x7)
- Per capita share of social insurance contributions (X8)
- Savings volume (x9)
- Number of private insurance funds (X10)
- Total Insurance Premiums (X11)
- Number of producers (X12)

Stepwise regression method was used to reach the best possible linear relationship between the explanatory variables and the dependent variable each time, and the results of applying the regression method were as follows:

First: the linear model

## 1) The regression equation for number of documents

By applying the normal linear regression, a linear relationship was found between the original values that express the number of insurance policies as a dependent variable and the group of independent variables used except for both the population and the individual's share of social insurance contributions, and the value of the corrected determination coefficient of the model was 0.43. But the significance tests showed that the model in its normal linear form has no statistical significance, whether with 95% or 99% confidence, and to verify this, the gradient regression method was used and the results of its application showed that there is no significant linear relationship between the dependent variable and the independent variables used.

2) The regression equation for insurance amounts

By applying the graduated linear regression, a strong linear relationship was found between the original values that express the insurance amounts as a dependent variable and a group of independent variables used as the value of the corrected determination coefficient of the model was 0.895. The regression equation was as follows:

 $\hat{Y}_{2} = -1260195 + 1.367X_{11} + 26084.098X_{9}$ 

The previous equation indicates the existence of a direct relationship between the amounts of insurance and each of the value of the insurance premiums in effect and the volume of savings, which is consistent with the assumptions of economic theory. The F-test was used to measure the significance of the regression model and the t-test to measure the significance of the regression equation coefficients. It was proved that the model as a whole has statistical significance and that the independent variables in the regression equation have a significant effect on the value of the dependent variable, whether with a 95% or 99% confidence degree.

Second: the semi-logarithmic model

1) The regression equation for number of documents

By applying the graduated semi-logarithmic regression model, it was concluded that there is a strong linear relationship between the logarithmic values of the number of insurance policies as a dependent variable and a group of independent variables used as the value of the corrected determination coefficient of the model was 0.995. The regression equation was as follows:

$$\ln(Y_1) = 7.579 + 1.071E - 6X_7 + 1.196E - 9X_{11} - 1.97E - 6X_{10}$$

The previous equation indicates that there is a direct relationship between the number of insurance policies and each of the size of the workforce in one million and the value of the applicable insurance premiums, and there is an inverse relationship between the number of insurance policies and the number of private insurance funds, which is consistent with the assumptions of economic theory. The F-test was used to measure the significance of the regression model and the t-test to measure the significance of the regression equation coefficients. It was proved that the model as a whole has statistical significance and that the independent variables in the regression equation have a significant effect on the value of the dependent variable, whether with a 95% or 99% confidence degree, except for the number of private insurance funds. Which affects statistically on the number of insurance policies at the level of 5% and does not affect the level of 1%.

2) The regression equation for insurance amounts

By applying the graded semi-logarithmic regression model, it was concluded that there is a strong linear relationship between the logarithmic values of the insurance amounts as a dependent variable and a group of independent variables used as the value of the corrected determination coefficient of the model was approximately 0.9995. The regression equation was as follows:

## $\ln(Y_2) = 10.054 + 2.038E - 5X_2 - 1.22E - 5X_7 - 2.37E - 3X_3 + 1.255E - 3X_5$

The previous equation indicates the existence of a direct relationship between insurance amounts and each of the population and per capita gross domestic product, and an inverse relationship between the amounts of insurance, the size of the workforce in millions, and the birth rate. The F-test was used to measure the significance of the regression model and the t-test to measure the significance of the regression equation coefficients. It was proved that the model as a whole has statistical significance and that the independent variables in the regression equation have a significant effect on the value of the dependent variable, either with a 95% or 99% confidence degree, except for the per capita share of the output. The gross domestic product, which statistically affects the amounts of insurance at the level of 5% and does not affect the level of 1%.

Third: the logarithmic model

#### 1) The regression equation for number of documents

By applying the logarithmic gradient regression model, it was concluded that there is a strong linear relationship between the logarithmic values of the number of insurance policies as a dependent variable and the logarithmic values of a group of independent variables used as the value of the corrected determination coefficient of the model was 0.993. The regression equation was as follows:

 $\ln(Y_1) = 7.411 + 1.872E - 2\ln(X_5) + 3.672E - 4\ln(X_{12})$ 

The previous equation indicates the existence of a direct relationship between the number of insurance policies and each of the per capita GDP and the number of insurance producers, which is consistent with the assumptions of economic theory. The F-test was used to measure the significance of the regression model and the t-test to measure the significance of the regression equation coefficients. It was proved that the model as a whole has statistical significance and that the independent variables in the regression equation have a significant effect on the value of the dependent variable, whether with a 95% or 99% confidence degree.

2) The regression equation for insurance amounts

By applying the graded logarithmic regression model, it was concluded that there is a strong linear relationship between the logarithmic values of the insurance amounts as a dependent variable and the logarithmic values of a group of independent variables used, as the value of the corrected determination coefficient of the model was approximately 0.976. The regression equation was as follows:

$$\ln(Y_{2}) = 1.702 + 0.939 \ln(X_{5}) - 2.07E - 2 \ln(X_{4})$$

The previous equation indicates that there is a direct relationship between the amounts of insurance and the per capita gross domestic product and that there is an inverse relationship between the amounts of insurance and the rates of marriage contracts. The F-test was used to measure the significance of the regression model and the t-test to measure the significance of the regression equation coefficients. It was proved that the model as a whole has statistical significance and that the independent variables in the regression equation have a significant effect on the value of the dependent variable with a 95% confidence degree, while at the 1% level of significance the rate of marriage contracts does not affect the amount of insurance.

The results of applying the regression method can be briefly summarized as follows:

a) Regression models for the number of insurance policies

Model	The coefficient of determina- tion	Explanatory variables
Linear	-	-
Half logarithmic	0.995	The size of the labor force, the premiums, the number of private funds
Logarithmic	0.993	Per capita output, the number of producers

It is evident from the previous table that the best regression models that can be used to describe the relationship between the number of insurance policies and the economic variables used in the study is the semi-logarithmic model, as it has the largest explanatory power (the value of the coefficient of determination = 0.995). It also includes a number of independent variables more than other models. Thus, the regression equation for the number of insurance policies is:

 $\ln(Y_1) = 7.579 + 1.071E - 6X_7 + 1.196E - 9X_{11} - 1.97E - 6X_{10}$ 

b) Regression models for insurance amounts

Model	The coefficient of determina- tion	Explanatory variables
Linear	0.895	Premiums, the amount of savings
Half logarithmic	0.9995	Population, size of the labor force, birth rates, per capita output
Logarithmic	0.976	Marriage contract rate, per capita output

It is evident from the previous table that the best regression models that can be used to describe the relationship between insurance amounts and economic variables used in the study is the semi-logarithmic model, as it has the largest explanatory power (the value of the coefficient of determination = 0.9995) and it also includes a number of independent variables more than other models. Thus, the equation for regression of insurance amounts is:

 $\ln(Y_2) = 10.054 + 2.038E - 5X_2 - 1.22E - 5X_7 - 2.37E - 3X_3 + 1.255E - 3X_5$ 

c) Property insurance and private insurance fund contributions

Curves matching was used, which includes the application of a set of linear and non-linear regression models to reach the best model in which the values of the dependent variable (public insurance premiums or private funds subscriptions) can be estimated that change almost regularly over time without being affected by the interlocking economic factors that govern demand The regression equations used are based on life insurance, and therefore the regression equations used are based on considering time as the independent variable Using the SPSS statistical program and data for the period from 1993 to 2006, the best model for estimating contributions was the following linear model:

C = 83479.7 + 99537(T)

Where C represents the dependent variable (subscriptions), and T represents the independent variable (time period).

The results also showed that the best model for estimating general insurance premiums is the following semi-logarithmic model:

Ln(PG) = 13.51 + 0.07445(T)

Where PG represents the dependent variable (direct underwriting premiums for total general insurance branches), and T represents the independent variable (time period). The summary of the assessment results based on the previous two models was as follows:

Year	Period	General insurance premiums	Private funds subscriptions
2005	15	2243954	1576535
2006	16	2417392	1676072
2007	17	2604236	1775609
2008	18	2805522	1875146
2009	19	3022364	1974683
2010	20	3255967	2074220

The estimated values according to the previous two models can be represented as an extension of the original peaks as follows:



## 4. Results of the study

The size of insurance amounts is related to a group of economic variables at the national level, including: population  $X_2$  - size of the labor force  $X_7$  - birth rates  $X_3$ , and it is also related to per capita GDP of  $X_5$ .

The number of insurance policies is related to a set of economic variables at the national level, including: the size of the work force X7 - the number of private insurance funds X10, and it is related to the total insurance premiums X11.

Diversifying efforts exerted to develop insurance awareness and market the service so that it is not limited to advertising mainly, but extends to developing serious plans in order to achieve increased demand for insurance service to protect individuals and property in addition to seeking to improve the image of the insurance sector and build a mutual culture between insurance companies and the public by consolidating and deepening thought Among the citizens, the mission of the insurance company is not only to collect premiums, but also to pay compensation to protect individuals and property.

### References

- Browne, Mark J., Jae Wook Chung, and Edward W. Frees (2020), "International Property-Liability Insurance Consumption", Journal of Risk and Insurance, vol. (67): 73-90. <u>https://doi.org/10.2307/253677</u>.
- Hoyt, R. and H. Kang (2020), "On the demand for corporate property insurance", Journal of Risk and Insurance, vol. (67): 91-107. https://doi.org/10.2307/253678.
- [3] Livio Di Matto and Herbert Emery (2020), "Self-Insurance and the Demand for Life Insurance : Private Evidence from Ontario", Lake head University, Ontario, ,available at: www.ucalgary.ca.
- [4] R. Narasimham (2020), "The Demographic trends Implications for Life Insurance in Asia-pacific", The Australian APEC Study Center, Melbourne, available at :www. Opec –finsecreg .org.
- [5] Loayza, N., Schmidt-Hebbel, K and L. Serven (2020), "What Drives Private Saving Across the World?", Review of Economics and Statistics, vol. (82): 165-181. <u>https://doi.org/10.1162/003465300558678</u>.
- [6] Joseph G. Eisenhauer (2022), "Relative Effects of premium Loading and Tax Deduction on the Demand for Insurance", Journal of Insurance Issues, available at: www.cob.fsu.edu.
- [7] Kari Gregory and Lance Elliot Kayfish (2022), "Genetic Testing in Life Insurance to be or not to be", Risk Management and Insurance, available at: www. ucalgary.ca.
- [8] Pablo Alonso and Irene Albarran Lozano (2022), "Modeling the Influence of the Purchasing Power Standard on the Insurance Policies Underwriting", University of Extermadura, Spain, available at: www.pascal.iseg.ut.pt.
- [9] Park, H., S.F. Borde and Y. Choi (2022), "Determinants of Insurance Pervasiveness: A Cross-National Analysis", International Business Review, vol. (11-1): 79-96 <u>https://doi.org/10.1016/S0969-5931(01)00048-8</u>.
- [10] Irene Albarran Lozano (2022), "Life Insurance ; An Empirical analysis Applied to Spain", University of Madrid, available in : www. ub . es .
- [11] Tienyu Hwanng and Simon Gao (2022),"The Determinants of the Demand for Life Insurance in an Emerging Economy the Case of China", University of Limerick ,Ireland, available at:www.mcu edu .tu.
- [12] Thorsten Beck and Ian Webb (2022), "Economic, Demographic and Institutional Determinants of Life Insurance Consumption Across Countries", World Bank and international Insurance Foundation, available at : www.worldbank.org.
- [13] Chee Chee Lim and Steven Haberman (2023), "Macroeconomics Variables and Demand for Life Insurance n Malaysia", Faculty of Actuarial Science and Statistics, CASS Business School, City University, London, available in:www. Staf.city.ac.uk.
- [14] Daniel Aunon and Nerin Paul Ehling (2023), "Corporate Demand for Property insurance", available Online.
- [15] Neil Esho (2023), "Law and the Demand for Property-Casualty Insurance Consumption", Center of International Economic Studies, University of Adelaide, Discussion Paper No. 0314.