

Using Multivariate Regression Analysis to Compute Longevity Risk and Actuarial Liability: The Case of Social Benefits for Workers in Venezuela

Evaristo Diz¹, J. Tim Query^{2,*}

¹Actuarial Science, Universidad Católica Andrés Bello, Caracas, Venezuela

²Finance Department, New Mexico State University, Caracas, Venezuela

*Corresponding author: tquery@nmsu.edu

Abstract In this study, a multivariate regression model is determined that allows computation of the actuarial liability of social benefits using a group of potential predictors. In general, the previous or independent predictors are the same as those utilized in an actuarial valuation of labor commitments. Several linear and non-linear models are considered and tested in this study. Among the most important findings of this research is that the PBO or Actuarial Liability depends fundamentally on a linear basis of two fundamental variables in the quantification of Social Benefits -- the Guarantees and the Social Benefits to Pay (PSP). The attractiveness of a relatively simple yet structurally robust model is a very attractive option in situations where managers are tasked to provide estimates in relatively short order, given that the two variables highlighted are generally readily available for most organizations. Realistically, companies are often required to give quick estimates, and they find it quite difficult to quantify an order of this magnitude. This model succeeds in filling this informational gap at the 4 percent interest rate.

Keywords: *multivariate analysis, actuarial liability, longevity*

Cite This Article: Evaristo Diz, and J. Tim Query, "Using Multivariate Regression Analysis to Compute Longevity Risk and Actuarial Liability: The Case of Social Benefits for Workers in Venezuela." *Journal of Finance and Economics*, vol. 6, no. 4 (2018): 144-153. doi: 10.12691/jfe-6-4-4.

1. Introduction

Longevity risk, as applied to pension funds or life insurance annuities, is the risk exposure of higher-than-expected payout ratios. Longevity risk exists due to the increasing life expectancy trends among policyholders and pensioners, and can result in payout levels that are higher than originally projected by a company. The types of plans exposed to the greatest levels of longevity risk are defined-benefit pension plans and annuities, which guarantee lifetime benefits for policy or plan holders.

Calculating future liability of social benefits is an increasingly important issue, in Venezuela and all over the world. The world's six largest pension saving systems – the United States, United Kingdom, Japan, Netherlands, Canada and Australia – are expecting a funding shortfall of \$224 trillion by 2050. If China and India are included, which have the world's largest populations, the combined savings gap for the eight countries reaches a total of \$400 trillion by 2050, a sum equivalent to five times the size of the current global economy [1].

The situation is especially dire in Venezuela. Venezuela's economy is built on oil—its leaders boast it has the world's largest proven reserves—and it is tempting to blame fickle crude prices for its woes. Venezuela has lost half of its economy since 2013. However, the problem

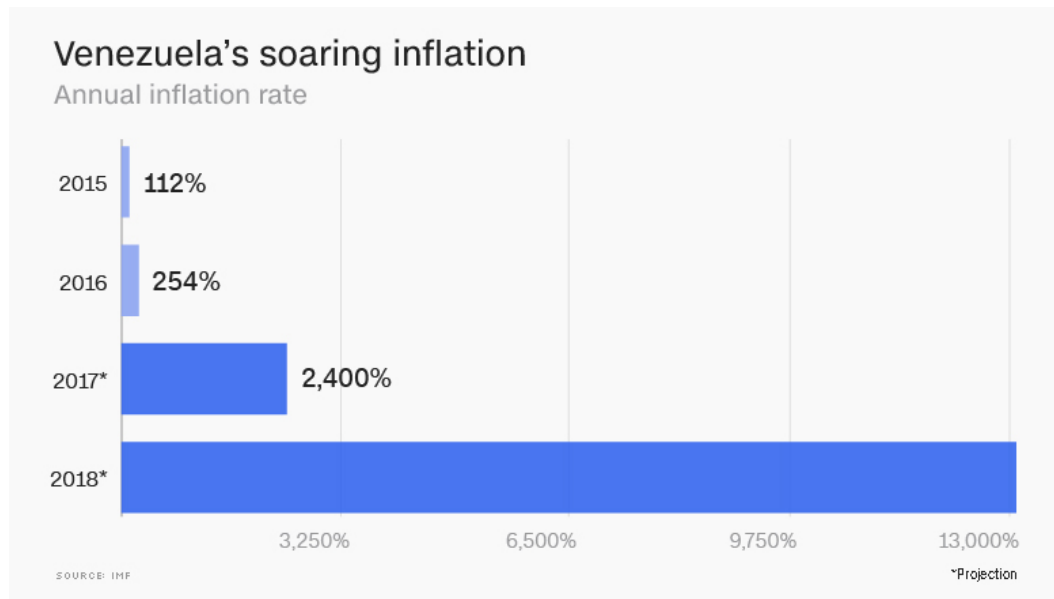
extends beyond oil and it's getting worse. Unemployment will reach 30% and prices on all types of goods in the country will rise 13,000% this year, according to figures published by the International Monetary Fund [2].

Pensions in Venezuela pay very little: 177,000 bolivares a month, or about \$5 on the black market, while a kilo of flour, rice or grain can easily cost up to 30,000 bolivares [3].

In this study, we determine a multivariate regression model that allows computation of the actuarial liability of social benefits using a group of potential predictors. In general, the previous or independent predictors are the same as those utilized in an actuarial valuation of labor commitments.

Creating a relevant and reasonably accurate Multivariate Regression Model would provide management timely estimates of actuarial liability levels by management for company employees entitled to Social Benefits. This would be a more efficient and economical alternative than engaging in a formal actuarial valuation. Obviously, as with any model, there is an estimation error at a predetermined confidence level. However, that error is not materially important in terms of the order of magnitude.

The rest of this paper is organized as follows. A brief review of relevant literature in the next section will be followed by a description of the methodology. A discussion of the results and their implications follows, with the conclusion closing the paper.



2. Literature Review

As pointed out by [4], the effort to model mortality dependence for multiple populations has important implications for mortality/longevity risk management. Their study presents a dynamic multi-population mortality analysis using copula models to assess multivariate dependence. Their simple and flexible, model in terms of model specification, is widely applicable to high dimension data.

The financial implications of ageing populations globally have garnered increased attention regarding the complex modeling of trends and volatility related to longevity risk. Financial models that allow for risk pricing and risk management have attracted increasing attention along with multiple factor models. A study by [5] investigates common trends through co-integration, and the factors driving the volatility of mortality using principal components analysis for many developed countries. They find the need for multiple factors when modeling mortality rates across all these countries, and that trends by countries are stochastic. They posit that the basic structure of the Lee-Carter model inadequately models the random variation and the full risk structure of mortality changes.

For organizations facing changes in a social benefit system that mandates employer contributions, a major problem facing decision-makers is forecasting the impact of those changes for companies in different sectors of the economy. The processing of information on an individualized basis can be nearly impossible when facing time constraints, making the development of a model imperative. [6] tackle the issues raised above by identifying key variables. The impact is explained using a reduced multivariable statistical model. This study uses a proprietary data set to model a methodology for measuring the rigor of these regulatory changes. Three key explanatory variables – the number of employees; the cumulative warranty; the monthly average full salary – are analyzed to determine the impact on those companies. Among these, the most important variable and in terms of statistical

significance is the amount of accumulated warranty at five percent, followed by the salary and the number of employees in the company, respectively.

Multivariate models are also used to value longevity risk premiums and can confirm that longevity-linked securities offer inexpensive opportunities for asset managers. [7] extends the standard Campbell-Viceira strategic asset allocation model by including a longevity-linked investment possibility. Their model estimation demonstrates that aggregate shocks to survival probabilities are reliable forecasters for long-term returns of longevity-linked securities.

3. Methodology

3.1. Description of the Multivariate Regression Model

To carry out the study's objectives, we begin with a proprietary database of 53 companies, comprised of a variety of firms representing different sizes, with diverse demographic variables and also with dissimilar liabilities, comparable to what we would expect to find in the real world. The filter for selection was that Pension Benefit Obligations (PBO) were determined at the same year of data, using 4 percent as the real rate of interest. From the previous date, a correlation analysis was performed to obtain the degree of association between the variables in order to rule out collinearity between them.

The variables used in the model are as follows:

1. Age: Average age of company employees
2. Guarantee: Social benefits accumulated by guarantee for the total employees
3. Number: Total number of employees
4. PBO (Pension Benefit Obligation): Actuarial Liability for Benefits
5. Benefits: Social benefits
6. Salary: Average Integral Salary of Employees
7. Service: Average Age

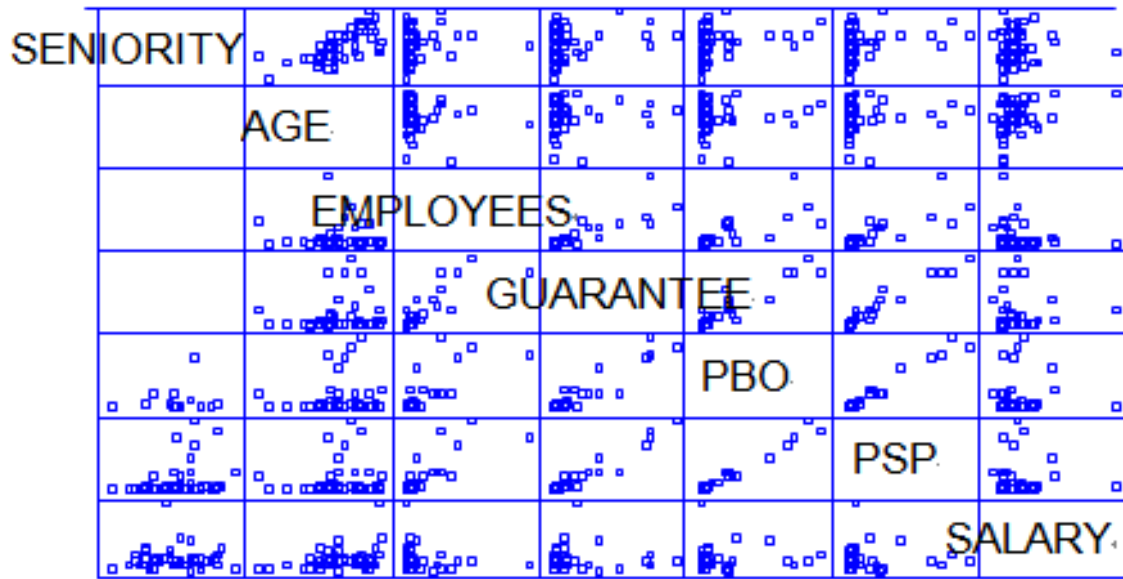


Chart 1. Correlation Matrix

Structure of the Data

$$\begin{matrix}
 Y & X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 \\
 \vdots & & & & & & & \\
 (\bar{Y}) & \dots\dots\dots & [X_{ij}] & \dots\dots\dots & & & & \\
 \vdots & & & & & & & \\
 \vdots & & & & & & &
 \end{matrix}$$

Vector Associated with PBO Dependent Variable
 Independent Variables Matrix.

3.2. Multivariate Analysis

Data / Variables:

- SENIORITY
- AGE
- EMPLOYEES
- WARRANTY
- PBO
- PSP (SOCIAL BENEFITS PAYABLE)
- SALARY

There are 53 complete cases used in the calculations. This procedure is designed to summarize several columns

of quantitative data. It will calculate several statistics, including correlations, covariances and partial correlations. Also included in the procedure are a series of multivariate graphs, which provide interesting observations of the data. From these results, a statistical model of the data will be constructed. The objective is to construct a model to predict the PBO given the values of one or more independent variables.

Table 1 below provides the statistical summary for each of the selected variables. It includes measures of central tendency, variability, and form. Of particular interest here is standardized bias and standardized kurtosis, which can be used to determine if the sample comes from a normal distribution.

3.3. Correlations

In Table 2 we can observe the Pearson product moment correlations, between each pair of variables. The range of these correlation coefficients ranges from -1 to +1, and they measure the strength of the linear relationship between the variables. The number of data pairings used to calculate each coefficient is also shown in brackets.

Table 1. Statistical Summary

| | SENIORITY | AGE | EMPLOYEES | GUARANTEE | PBO | PSP | SALARY |
|--------------------------|-----------|----------|-----------|-----------|-----------|-----------|----------|
| Number | 53 | 53 | 53 | 53 | 53 | 53 | 53 |
| Average | 7,85472 | 38,6547 | 543,679 | 2,26399E8 | 3,03923E8 | 7,30446E8 | 180369, |
| Geometric Mean | 7,28962 | 38,4643 | 163,89 | 7,69466E7 | 7,13146E7 | 2,05604E8 | 140313, |
| Standard Deviation | 2,79915 | 3,74657 | 1060,34 | 3,78329E8 | 5,59792E8 | 1,34296E9 | 149201, |
| Coefficient of Variation | 35,6365% | 9,69239% | 195,031% | 167,107% | 184,189% | 183,854% | 82,7203% |
| Mínimum | 1,5 | 27,2 | 3,0 | 532978, | 46828,0 | 632149, | 3608,0 |
| Máximum | 14,1 | 44,7 | 6179,0 | 1,66448E9 | 2,41622E9 | 6,06969E9 | 990681, |
| Range | 12,6 | 17,5 | 6176,0 | 1,66395E9 | 2,41617E9 | 6,06906E9 | 987073, |
| Standardized Bias | 0,226984 | -2,3426 | 10,6964 | 7,4198 | 7,74638 | 8,17609 | 10,1983 |
| Standardized Curtosis | -0,883473 | 1,8201 | 23,128 | 8,38428 | 9,09612 | 10,5993 | 24,4371 |

Table 2. Correlation

| | SENIORITY | AGE | EMPLOYEES | GUARANTEE | PBO | PSP | SALARY |
|-----------|-----------|---------|-----------|-----------|--------|--------|---------|
| SENIORITY | | 0,6665 | 0,1979 | 0,3531 | 0,3354 | 0,3350 | 0,0120 |
| | | (53) | (53) | (53) | (53) | (53) | (53) |
| | | 0,0000 | 0,1554 | 0,0095 | 0,0141 | 0,0142 | 0,9321 |
| AGE | 0,6665 | | -0,0891 | 0,1079 | 0,0915 | 0,0982 | 0,3571 |
| | (53) | | (53) | (53) | (53) | (53) | (53) |
| | 0,0000 | | 0,5256 | 0,4420 | 0,5146 | 0,4843 | 0,0087 |
| EMPLOYEES | 0,1979 | -0,0891 | | 0,8522 | 0,7860 | 0,8275 | -0,1769 |
| | (53) | (53) | | (53) | (53) | (53) | (53) |
| | 0,1554 | 0,5256 | | 0,0000 | 0,0000 | 0,0000 | 0,2051 |
| GUARANTEE | 0,3531 | 0,1079 | 0,8522 | | 0,9398 | 0,9765 | 0,0297 |
| | (53) | (53) | (53) | | (53) | (53) | (53) |
| | 0,0095 | 0,4420 | 0,0000 | | 0,0000 | 0,0000 | 0,8326 |
| PBO | 0,3354 | 0,0915 | 0,7860 | 0,9398 | | 0,9850 | 0,0562 |
| | (53) | (53) | (53) | (53) | | (53) | (53) |
| | 0,0141 | 0,5146 | 0,0000 | 0,0000 | | 0,0000 | 0,6895 |
| PSP | 0,3350 | 0,0982 | 0,8275 | 0,9765 | 0,9850 | | 0,0311 |
| | (53) | (53) | (53) | (53) | (53) | | (53) |
| | 0,0142 | 0,4843 | 0,0000 | 0,0000 | 0,0000 | | 0,8251 |
| SALARY | 0,0120 | 0,3571 | -0,1769 | 0,0297 | 0,0562 | 0,0311 | |
| | (53) | (53) | (53) | (53) | (53) | (53) | |
| | 0,9321 | 0,0087 | 0,2051 | 0,8326 | 0,6895 | 0,8251 | |

The third number in each block of the table is a P-value that tests the statistical significance of the estimated correlations. P-values below 0.05 indicate correlations significantly different from zero, with a confidence level of 95.0%. The following pairs of variables have P-values below 0.05: SENORITY and AGE

SENIORITY and GUARANTEE

SENIORITY and PBO

SENIORITY and PSP

AGE and SALARY

EMPLOYEES AND GUARANTEE

EMPLOYEES and PBO

EMPLOYEES and PSP

GUARANTEE and PBO

GUARANTEE and PSP PBO and SP

3.4. Selection of Regression Model - PBO

In order to select the best multivariate model that explains the association of PBO with one or more predictor variables, the different combinations of the dependent variable are run with all the independent variables.

Dependent variable: PBO

Independent variables:

A = SENIORITY

B = AGE

C = EMPLOYEES

D = WARRANTY

E = PSP

F = SALARY

Number of completed cases: 53

Number of models adjusted: 63

The results of fitting multiple regression models to describe the relationship between PBO and 6 predictor variables is illustrated in Table 3. The models have been adjusted containing all combinations from 0 to 5 variables. Tabulated statistics include mean error square (CME), adjusted and unadjusted R-Square values, and the Mallows Cp statistic.¹ To determine which models are best according to these different criteria.

4. Results

In this section, we present and analyze our results. The results of fitting multiple regression models to describe the relationship between PBO and 6 predictor variables is illustrated in Table 3. The models have been adjusted containing all combinations from 0 to 5 variables. Tabulated statistics include mean error square, adjusted and unadjusted R-Square values, and the Mallows Cp statistic. These measures are used to help determine which models are best according to these different criteria.

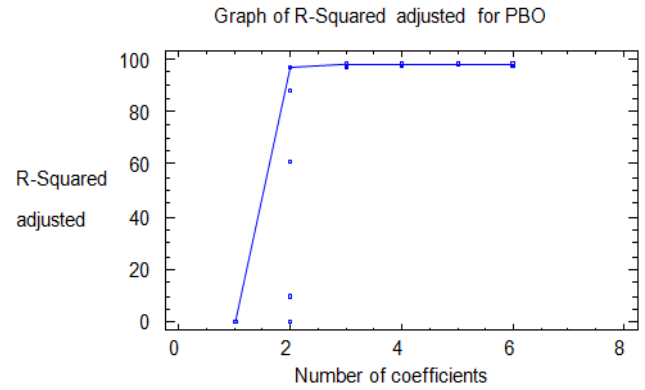
Table 4 displays the models that give the highest values of adjusted R-squared. The adjusted R-squared statistic measures the proportion of variability in PBO that is explained by the model. Large values of adjusted R-squares correspond to small mean square error values. Up to 5 models are shown in each subset of 0 to 5 variables.

¹Mallow's Cp is used to assess the fit of a regression model that has been estimated using ordinary least squares. It is applied in the context of model selection, where many predictor variables are available for predicting some outcome, and the goal is to find the best model involving a subset of these predictors. A small value of Cp means that the model is relatively precise.

The best model contains 5 variables, SENIORITY, AGE, GUARANTEE, PSP, and SALARY.

Table 3. Results of the model

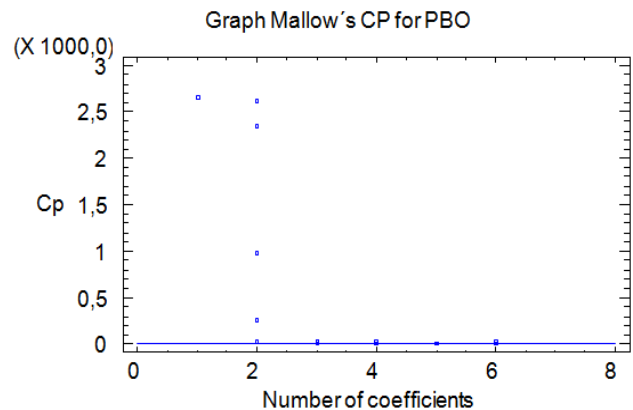
| Mean Square Error | R-Squared | Adjusted R-Squared | Mallow's Cp | Included Variables |
|-------------------|-----------|--------------------|-------------|--------------------|
| 3,13367E17 | 0,0 | 0,0 | 2643,84 | |
| 2,83574E17 | 11,2476 | 9,50731 | 2342,74 | A |
| 3,13367E17 | 1,92308 | 0,0 | 2623,28 | B |
| 1,22115E17 | 61,7807 | 61,0313 | 980,948 | C |
| 3,73013E16 | 88,3255 | 88,0966 | 265,609 | D |
| 9,49002E15 | 97,0298 | 96,9716 | 31,0412 | E |
| 3,13367E17 | 1,92308 | 0,0 | 2637,34 | F |
| 2,79027E17 | 14,3831 | 10,9584 | 2260,24 | AB |
| 1,13591E17 | 65,1457 | 63,7515 | 892,269 | AC |
| 3,80428E16 | 88,3269 | 87,86 | 267,571 | AD |
| 9,66914E15 | 97,0331 | 96,9144 | 32,9529 | AE |
| 2,88359E17 | 11,5195 | 7,98033 | 2337,41 | AF |
| 1,15981E17 | 64,4124 | 62,9889 | 912,029 | BC |
| 3,80152E16 | 88,3354 | 87,8688 | 267,343 | BD |
| 9,67089E15 | 97,0326 | 96,9139 | 32,9674 | BE |
| 3,13367E17 | 3,84615 | 0,0 | 2623,57 | BF |
| 3,77822E16 | 88,4069 | 87,9431 | 265,416 | CD |
| 8,80233E15 | 97,2991 | 97,191 | 25,7853 | CE |
| 1,11735E17 | 65,715 | 64,3436 | 876,926 | CF |
| 6,25344E15 | 98,0812 | 98,0044 | 4,70894 | DE |
| 3,77876E16 | 88,4052 | 87,9414 | 265,461 | DF |
| 9,46706E15 | 97,0951 | 96,9789 | 31,2819 | EF |
| 1,15168E17 | 65,3686 | 63,2483 | 888,261 | ABC |
| 3,87243E16 | 88,3554 | 87,6425 | 268,802 | ABD |
| 9,81574E15 | 97,0484 | 96,8677 | 34,5418 | ABE |
| 2,76947E17 | 16,7208 | 11,622 | 2199,24 | ABF |
| 3,85515E16 | 88,4074 | 87,6977 | 267,402 | ACD |
| 8,98053E15 | 97,2995 | 97,1342 | 27,7737 | ACE |
| 1,03949E17 | 68,742 | 66,8282 | 797,354 | ACF |
| 6,26069E15 | 98,1174 | 98,0021 | 5,73347 | ADE |
| 3,85543E16 | 88,4066 | 87,6968 | 267,424 | ADF |
| 9,64952E15 | 97,0983 | 96,9207 | 33,1948 | AEF |
| 3,84039E16 | 88,4518 | 87,7448 | 266,206 | BCD |
| 8,81933E15 | 97,348 | 97,1856 | 26,4674 | BCE |
| 1,10732E17 | 66,7025 | 64,6639 | 852,314 | BCF |
| 6,38098E15 | 98,0812 | 97,9637 | 6,70826 | BDE |
| 3,84065E16 | 88,451 | 87,7439 | 266,227 | BDF |
| 9,58205E15 | 97,1186 | 96,9422 | 32,6481 | BEF |
| 6,30064E15 | 98,1054 | 97,9894 | 6,05718 | CDE |
| 3,84376E16 | 88,4417 | 87,734 | 266,479 | CDF |
| 8,96421E15 | 97,3044 | 97,1394 | 27,6415 | CEF |
| 6,16898E15 | 98,145 | 98,0314 | 4,99033 | DEF |
| 3,91236E16 | 88,4755 | 87,5151 | 267,568 | ABCD |
| 8,90361E15 | 97,3773 | 97,1587 | 27,6779 | ABCE |
| 1,05558E17 | 68,9061 | 66,315 | 794,93 | ABCF |
| 6,2961E15 | 98,1454 | 97,9908 | 6,97917 | ABDE |
| 3,89466E16 | 88,5276 | 87,5716 | 266,163 | ABDF |
| 9,58214E15 | 97,1774 | 96,9422 | 33,0642 | ABEF |
| 6,34409E15 | 98,1312 | 97,9755 | 7,36016 | ACDE |
| 3,92382E16 | 88,4417 | 87,4785 | 268,478 | ACDF |
| 9,15003E15 | 97,3047 | 97,0801 | 29,634 | ACEF |
| 6,17529E15 | 98,181 | 98,0294 | 6,02015 | ADEF |
| 6,42246E15 | 98,1082 | 97,9505 | 7,98223 | BCDE |
| 3,89933E16 | 88,5139 | 87,5567 | 266,533 | BCDF |
| 8,93677E15 | 97,3675 | 97,1481 | 27,9411 | BCEF |
| 6,26937E15 | 98,1532 | 97,9994 | 6,76701 | BDEF |
| 6,28414E15 | 98,1489 | 97,9946 | 6,88425 | CDEF |
| 6,32653E15 | 98,1752 | 97,9811 | 8,17451 | ABCDE |
| 3,95915E16 | 88,5806 | 87,3658 | 266,734 | ABCDF |
| 8,93983E15 | 97,4215 | 97,1472 | 28,487 | ABCEF |
| 5,93028E15 | 98,2895 | 98,1076 | 5,0945 | ABDEF |
| 6,30534E15 | 98,1813 | 97,9879 | 8,00979 | ACDEF |
| 6,376E15 | 98,161 | 97,9653 | 8,55901 | BCDEF |



Graphic 2. Adjusted R2 for pension benefit obligation

Table 4. The models with higher Adjusted R-Square

| CME | R-Square | Adjusted R-Square | Cp | Variables Included |
|------------|----------|-------------------|---------|--------------------|
| 5,93028E15 | 98,2895 | 98,1076 | 5,0945 | ABDEF |
| 6,16898E15 | 98,145 | 98,0314 | 4,99033 | DEF |
| 6,17529E15 | 98,181 | 98,0294 | 6,02015 | ADEF |
| 6,25344E15 | 98,0812 | 98,0044 | 4,70894 | DE |
| 6,26069E15 | 98,1174 | 98,0021 | 5,73347 | ADE |
| 6,26937E15 | 98,1532 | 97,9994 | 6,76701 | BDEF |
| 6,28414E15 | 98,1489 | 97,9946 | 6,88425 | CDEF |
| 6,2961E15 | 98,1454 | 97,9908 | 6,97917 | ABDE |
| 6,30064E15 | 98,1054 | 97,9894 | 6,05718 | CDE |
| 6,30534E15 | 98,1813 | 97,9879 | 8,00979 | ACDEF |
| 6,32653E15 | 98,1752 | 97,9811 | 8,17451 | ABCDE |
| 6,34409E15 | 98,1312 | 97,9755 | 7,36016 | ACDE |
| 6,376E15 | 98,161 | 97,9653 | 8,55901 | BCDEF |
| 6,38098E15 | 98,0812 | 97,9637 | 6,70826 | BDE |
| 8,80233E15 | 97,2991 | 97,191 | 25,7853 | CE |
| 8,81933E15 | 97,348 | 97,1856 | 26,4674 | BCE |
| 8,93983E15 | 97,4215 | 97,1472 | 28,487 | ABCEF |
| 9,46706E15 | 97,0951 | 96,9789 | 31,2819 | EF |
| 9,49002E15 | 97,0298 | 96,9716 | 31,0412 | E |
| 9,66914E15 | 97,0331 | 96,9144 | 32,9529 | AE |
| 9,67089E15 | 97,0326 | 96,9139 | 32,9674 | BE |
| 3,73013E16 | 88,3255 | 88,0966 | 265,609 | D |
| 1,22115E17 | 61,7807 | 61,0313 | 980,948 | C |
| 2,83574E17 | 11,2476 | 9,50731 | 2342,74 | A |
| 3,13367E17 | 1,92308 | 0,0 | 2637,34 | F |
| 3,13367E17 | 0,0 | 0,0 | 2643,84 | |



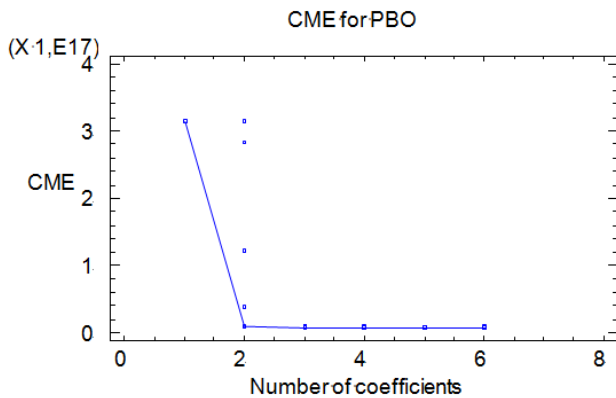
Graphic 3. Mallow's Cp for the benefit of pension Obligations

Table 5 below shows the models that give the smaller values of the statistical Cp of Mallows. Cp is a measure of

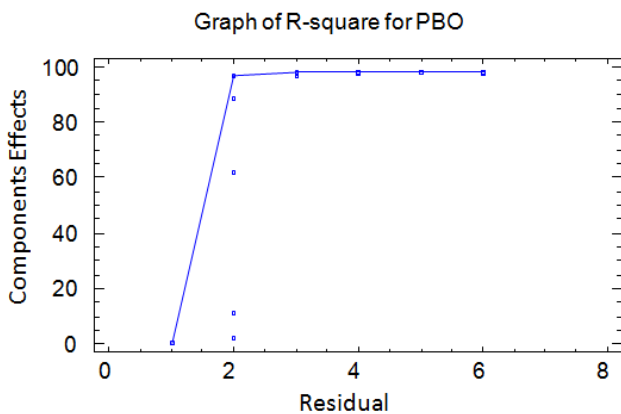
bias in the model, based on the comparison between the mean square of the total error and the variance of the true error. Non-bias models have an expected value of about p , where p is the number of coefficients in the adjusted model (including the constant). You should look for models with C_p values close to p .

Table 5. The models with lower C_p

| <i>CME</i> | <i>R-Square</i> | <i>Adjusted R-Square</i> | <i>Cp</i> | <i>Variables Included</i> |
|------------|-----------------|--------------------------|-----------|---------------------------|
| 6,25344E15 | 98,0812 | 98,0044 | 4,70894 | DE |
| 6,16898E15 | 98,145 | 98,0314 | 4,99033 | DEF |
| 5,93028E15 | 98,2895 | 98,1076 | 5,0945 | ABDEF |
| 6,26069E15 | 98,1174 | 98,0021 | 5,73347 | ADE |
| 6,17529E15 | 98,181 | 98,0294 | 6,02015 | ADEF |
| 6,30064E15 | 98,1054 | 97,9894 | 6,05718 | CDE |
| 6,38098E15 | 98,0812 | 97,9637 | 6,70826 | BDE |
| 6,26937E15 | 98,1532 | 97,9994 | 6,76701 | BDEF |
| 6,28414E15 | 98,1489 | 97,9946 | 6,88425 | CDEF |
| 6,2961E15 | 98,1454 | 97,9908 | 6,97917 | ABDE |
| 6,34409E15 | 98,1312 | 97,9755 | 7,36016 | ACDE |
| 6,30534E15 | 98,1813 | 97,9879 | 8,00979 | ACDEF |
| 6,32653E15 | 98,1752 | 97,9811 | 8,17451 | ABCDE |
| 6,376E15 | 98,161 | 97,9653 | 8,55901 | BCDEF |
| 8,80233E15 | 97,2991 | 97,191 | 25,7853 | CE |
| 8,81933E15 | 97,348 | 97,1856 | 26,4674 | BCE |
| 8,93983E15 | 97,4215 | 97,1472 | 28,487 | ABCEF |
| 9,49002E15 | 97,0298 | 96,9716 | 31,0412 | E |
| 9,46706E15 | 97,0951 | 96,9789 | 31,2819 | EF |
| 9,66914E15 | 97,0331 | 96,9144 | 32,9529 | AE |
| 9,67089E15 | 97,0326 | 96,9139 | 32,9674 | BE |
| 3,73013E16 | 88,3255 | 88,0966 | 265,609 | D |
| 1,22115E17 | 61,7807 | 61,0313 | 980,948 | C |
| 2,83574E17 | 11,2476 | 9,50731 | 2342,74 | A |
| 3,13367E17 | 1,92308 | 0,0 | 2623,28 | B |
| 3,13367E17 | 0,0 | 0,0 | 2643,84 | |



Graphic 4. The square of the average errors for the benefit of pension obligations



Graphic 5. R-square for pension commitments

4.1. Multiple Regression - PBO

Based on the previous selection we proceeded to use the model with two independent variables having the lower C_p and also utilized the principle of parsimony regarding the smallest number of possible variables.

Dependent variable: PBO

Independent variables:

PSP

WARRANTY

| <i>Parameter</i> | <i>Estimation</i> | <i>Standard Error</i> | <i>Statistical T</i> | <i>P-Value</i> |
|------------------|-------------------|-----------------------|----------------------|----------------|
| CONSTANT | 2,19292E7 | 1,28568E7 | 1,70565 | 0,0943 |
| PSP | 0,60433 | 0,0379033 | 15,944 | 0,0000 |
| GUARANTEE | -0,704226 | 0,134545 | -5,23411 | 0,0000 |

5.1. Analysis of Variance

| Source | Sum of Squares | GI | Mid Square | F-Ratio | P-Value |
|---------------|----------------|----|------------|---------|---------|
| Model | 1,59824E19 | 2 | 7,99121E18 | 1277,89 | 0,0000 |
| Residual | 3,12672E17 | 50 | 6,25344E15 | | |
| Total (Corr.) | 1,62951E19 | 52 | | | |

R-square = 98.0812 percent

R-square (adjusted for g.l.) = 98.0044 percent

Standard error of est. = 7.90787E7

Mean Absolute Error = 3.66167E7

Durbin-Watson Statistic = 1.97114 (P = 0.4729)

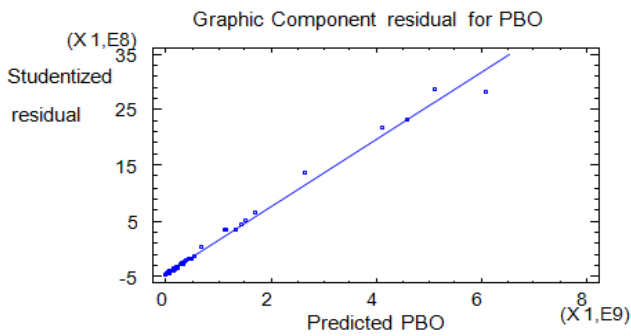
Autocorrelation of residuals in delay 1 = 0.0136931

The output shows the results of fitting a multiple linear regression model to describe the relationship between PBO and 2 independent variables. The equation of the fitted model is

$$PBO = 2.19292E7 + 0.60433 * PSP - 0.704226 * WARRANTY.$$

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables with a confidence level of 95.0%. The R-Squared statistic indicates that the adjusted model explains 98.0812% of the variability in PBO. The adjusted R-Squared statistic, which is more appropriate to compare models with different number of independent variables, is 98.0044%. The standard error of the estimate shows that the standard deviation of the residuals is 7.90787E7. This value can be used to construct limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 3.66167E7 is the average value of the waste. The Durbin-Watson (DW) statistic examines the residuals to determine if there is any significant correlation based on the order in which they are presented in the data file. Since the P-value is greater than 0.05, there is no indication of a serial autocorrelation in the residues with a confidence level of 95.0%.

To determine if the model can be simplified, note that the highest P-value of the independent variables is 0.0000, which corresponds to GUARANTEE. Since the P-value is less than 0.05, that term is statistically significant with a confidence level of 95.0%. Consequently, we probably should not remove any variables from the model.

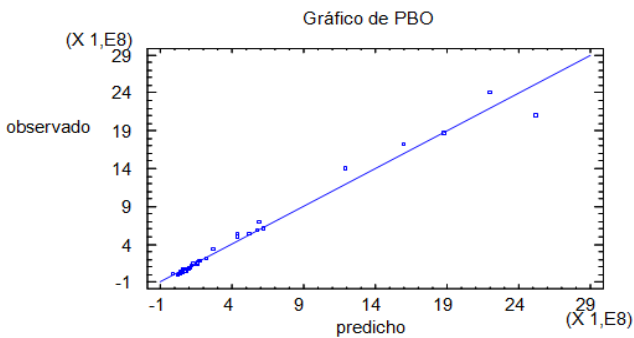


Graph 7. Graphic Component Residual

5.2. Additional ANOVA for Variables in the Adjusted Order

| Source | Sum of Squares | Gl | Medium Squared | F - Reason | P-Value |
|-----------|----------------|----|----------------|------------|---------|
| PSP | 1,58111E19 | 1 | 1,58111E19 | 2528,38 | 0,0000 |
| GUARANTEE | 1,71319E17 | 1 | 1,71319E17 | 27,40 | 0,0000 |
| Model | 1,59824E19 | 2 | | | |

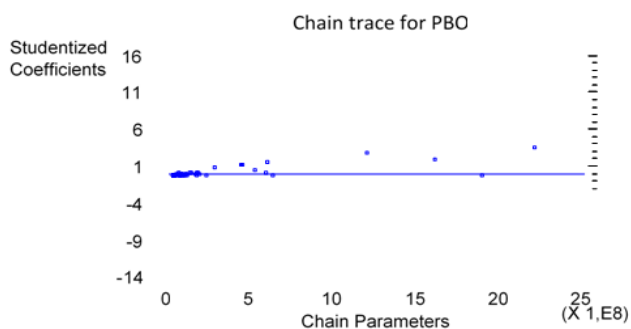
This table shows the statistical significance of each variable as it was added to the model. You can use this table to help you determine if the model can be simplified, especially if you are adjusting a polynomial. In sample how the relation is linear or polynomial of degree 1.



Graph 8. Graph of PBO

5.3. 95.0 Percent Confidence Intervals for Coefficient Estimates

| Parameter | Estimation | Standard Error | Lower Limit | Upper Limit |
|-----------|------------|----------------|-------------|-------------|
| CONSTANT | 2,19292E7 | 1,28568E7 | -3,89442E6 | 4,77528E7 |
| PSP | 0,60433 | 0,0379033 | 0,528198 | 0,680461 |
| GUARANTEE | -0,704226 | 0,134545 | -0,974469 | -0,433983 |



Graph 9. Chain Trace

This table shows 95.0% confidence intervals for the coefficients in the model. The confidence intervals show how accurately the coefficients can be estimated given the amount of data available, and the level of noise that is present.

Table 7. Matrix of correlation coefficients

| | CONSTANT | PSP | GUARANTEE |
|----------|----------|---------|-----------|
| CONSTANT | 1,0000 | 0,1602 | -0,2664 |
| PSP | 0,1602 | 1,0000 | -0,9765 |
| GURANTEE | -0,2664 | -0,9765 | 1,0000 |

This table shows the estimated correlations between the coefficients in the adjusted model. These correlations are used to detect the presence of severe multicollinearity, that is, correlation between the predictor variables. In this case, there is 1 correlation with absolute value greater than 0.5 (not including the constant term).

5.4. Atypical Residues

This table contains information about PBO, which is generated using the adjusted model. The table includes:

- (1) predicted PBO values using the adjusted model
- (2) the standard error for each predicted value
- (3) 95.0% forecast intervals for new observations
- (4) 95.0% confidence intervals for the mean response

Table 8. Atypical Waste

| Row | Y | Predict | Residual | Studentized |
|-----|-----------|-----------|------------|-------------|
| 9 | 2,41622E9 | 2,20013E9 | 2,16086E8 | 3,56 |
| 32 | 2,10933E9 | 2,51785E9 | -4,08522E8 | -13,64 |
| 40 | 1,39299E9 | 1,19285E9 | 2,00142E8 | 2,89 |

The table of atypical residues lists all the observations that have Residues greater than 2, in absolute value. The Remaining residues measure how many standard deviations deviate each observed PBO value from the adjusted model, using all data except that observation. In this case, there are 3 researched residues greater than 2, 2 greater than 3. It is convenient to examine the observations with residues greater than 3 to determine if they are aberrant values that should be removed from the model and treated separately. In fact, only 2 out of 53 observations would be within the classification of atypical residues in our literal opinion and the fundamental reason is that these two companies are significantly different from the rest.

5.5. Chain Regression - PBO

This procedure is designed to provide estimates of regression coefficients when independent variables are strongly correlated. By allowing a slight amount of bias (bias), the accuracy of the estimates can often increase greatly. In this case, the adjusted regression model is

Dependent variable: PBO

Independent variables:

PSP

WARRANTY

Number of completed cases: 53

Table 9. The results of the model for the parameter string = 0.0

| Parameter | Estimated | Factor of Inflation of Variance |
|-----------|-----------|---------------------------------|
| CONSTANTE | 2,19292E7 | |
| PSP | 0,60433 | 21,5458 |
| GARANTIA | -0,704226 | 21,5458 |

R-Square = 98.0812 percent
 R-Square (adjusted by g.l.) = 98.0044 percent
 Standard error of est. = 7.90787E7
 Absolute mean error = 3.66167E7
 Durbin-Watson Statistic = 1.97114
 Residual delay autocorrelation 1 = 0.0136931.

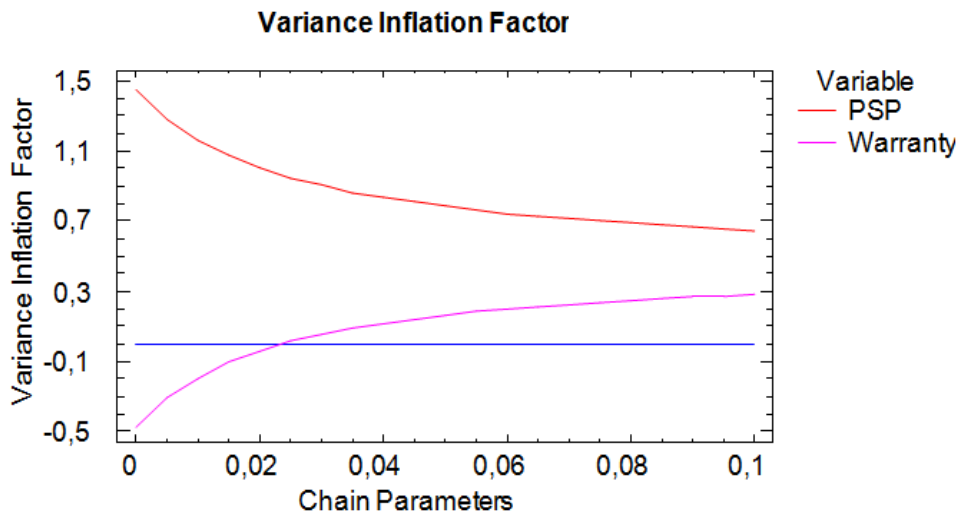
Table 10. Residual Analysis

| | Estimation |
|------|------------|
| n | 53 |
| CME | 6,25344E15 |
| MAE | 3,66167E7 |
| MAPE | 1715,33 |
| ME | 4,49846E-9 |
| MPE | -1703,66 |

$$PBO = 2.19292E7 + 0.60433 * PSP - 0.704226 * WARRANTY.$$

The current value of the string parameter is 0.0, which is equivalent to ordinary least squares. The string parameter is normally set between 0 and 1. To determine a good value for the string parameter, you should examine the standardized regression coefficients or the variance inflation factors.

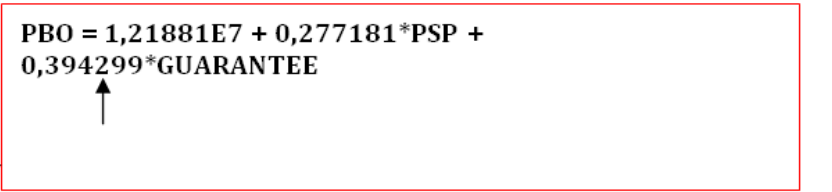
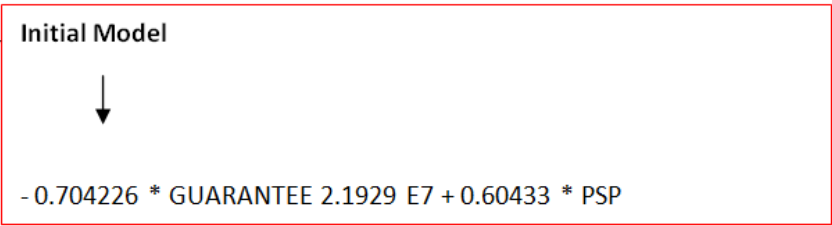
The R-Squared statistic indicates that the model, thus adjusted, explains 98.0812% of the variability in PBO. The adjusted R-Squared statistic, which is more suitable for comparing models with different number of independent variables, is 98.0044%. The standard error of the estimate shows that the standard deviation of the residuals is 7.90787E7. The mean absolute error (MAE) of 3.66167E7 is the average value of the waste. The Durbin-Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they were presented in their data file.



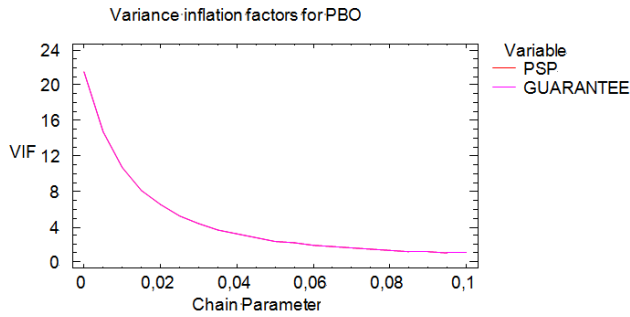
Graph 10.

Table 11. The regression coefficients

| String | PSP | GUARANTEE |
|--------|----------|------------|
| 0,0 | 0,60433 | -0,704226 |
| 0,005 | 0,533359 | -0,455938 |
| 0,01 | 0,483435 | -0,28234 |
| 0,015 | 0,446354 | -0,154314 |
| 0,02 | 0,417687 | -0,0561356 |
| 0,025 | 0,394832 | 0,0214316 |
| 0,03 | 0,376158 | 0,0841711 |
| 0,035 | 0,360595 | 0,135888 |
| 0,04 | 0,347407 | 0,17919 |
| 0,045 | 0,336076 | 0,21592 |
| 0,05 | 0,326221 | 0,247423 |
| 0,055 | 0,317563 | 0,274699 |
| 0,06 | 0,309886 | 0,298509 |
| 0,065 | 0,303024 | 0,319442 |
| 0,07 | 0,296846 | 0,337961 |
| 0,075 | 0,29125 | 0,354435 |
| 0,08 | 0,286151 | 0,369161 |
| 0,085 | 0,28148 | 0,382383 |
| 0,09 | 0,277181 | 0,394299 |
| 0,095 | 0,273208 | 0,405077 |
| 0,1 | 0,26952 | 0,414855 |



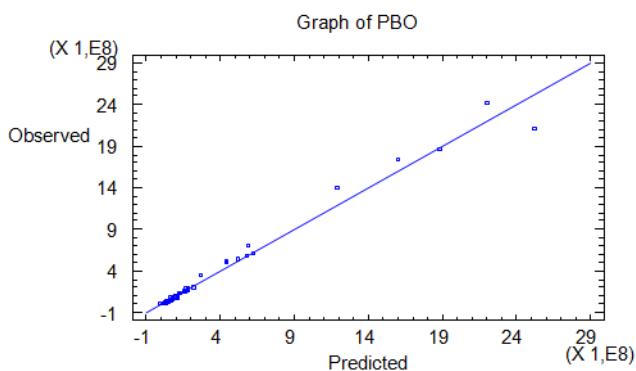
This table shows the estimated regression coefficients for chain parameter values between 0.0 and 0.1. As the string parameter is incremented from 0, the coefficients often change dramatically at first, but then become relatively stable. A good value for the string parameter is the smallest value after which the estimates change slowly.



Graphic 11.

Table 12. Standardized regression coefficients

| String | PSP | Guarantee |
|--------|----------|------------|
| 0,0 | 1,4498 | -0,475943 |
| 0,005 | 1,27954 | -0,308141 |
| 0,01 | 1,15977 | -0,190816 |
| 0,015 | 1,07082 | -0,104291 |
| 0,02 | 1,00204 | -0,0379386 |
| 0,025 | 0,947213 | 0,0144843 |
| 0,03 | 0,902415 | 0,0568861 |
| 0,035 | 0,865078 | 0,0918385 |
| 0,04 | 0,83344 | 0,121103 |
| 0,045 | 0,806255 | 0,145927 |
| 0,05 | 0,782615 | 0,167218 |
| 0,055 | 0,761843 | 0,185652 |
| 0,06 | 0,743425 | 0,201744 |
| 0,065 | 0,726963 | 0,215892 |
| 0,07 | 0,712143 | 0,228407 |
| 0,075 | 0,698717 | 0,239541 |
| 0,08 | 0,686484 | 0,249494 |
| 0,085 | 0,675278 | 0,258429 |
| 0,09 | 0,664965 | 0,266483 |
| 0,095 | 0,655433 | 0,273767 |
| 0,1 | 0,646587 | 0,280375 |



Graph 12.

Table 12 shows the standardized regression coefficients for regression parameter values between 0.0 and 0.1. These are the coefficients of the regression model when the variables are expressed in standardized form. As the

string parameter is incremented from 0, the coefficients often change dramatically at first, but then become relatively stable. A good value for the string parameter is the smallest value after which the estimates change slowly. It is recognized that this is very subjective, but the chain trace can help you make a good choice.

The model finally chosen is:

Results of the Model for the Chain Parameter = 0.09

| Parameter | Estimate | Factor of Variance Inflation |
|-----------|-----------|------------------------------|
| CONSTANT | 1,21881E7 | |
| PSP | 0,277181 | 1,14312 |
| GUARANTEE | 0,394299 | 1,14312 |

R-Square = 90.546 percent

R-Square (adjusted by g.l.) = 90.1679 percent

Standard error of est. = 1.25531 E8

Absolute mean error = 6.29326E7

Durbin-Watson Statistic = 2.22691

Residual delay autocorrelation 1 = -0.1113716

Table 13. Residual analysis

| | Estimation | Validation |
|------|------------|------------|
| n | 53 | |
| CME | 1,57581E16 | |
| MAE | 6,29326E7 | |
| MAPE | 1083,41 | |
| ME | 7,31E-9 | |
| MPE | -1075,06 | |

This model is designed to provide estimates of regression coefficients when independent variables are strongly correlated. By allowing a slight amount of bias (bias), the accuracy of the estimates can often be greatly increased. In this case, the adjusted regression model is

$$PBO = 1.21881E7 + 0.277181 * PSP + 0.394299 * WARRANTY.$$

The current value of the string parameter is 0.09. The string parameter is normally set between 0 and 1. To determine a good value for the string parameter, you should examine the standardized regression coefficients or the variance inflation factors. Generally, this model unlike the previous one, could be more stable in predictions even beyond the domain of the data.

The R-Squared statistic indicates that the model, thus adjusted, explains 90.546% of the variability in PBO. The adjusted R-Squared statistic, which is more adequate to compare models with different number of independent variables, is 90.1699%. The standard error of the estimate shows that the standard deviation of the residues is 1.25531E8. The mean absolute error (MAE) of 6.29326E7 is the average value of the waste.

In many cases, the estimates of the coefficients will change dramatically as it moves away from 0. Eventually, the coefficients will stabilize as they do in the neighborhood of = 0.09. Unfortunately, they no longer apply statistical significance tests, so it is difficult to determine which coefficients are statistically significant. However, it is likely that the highlighting model is much more stable than the ordinary least squares model,

particularly when extrapolated outside the experimental region.

6. Conclusion

It is important to note that Actuarial Liabilities, as is well known, depend on a series of financial and demographic assumptions. Several linear and non-linear models are considered and tested in this study. Among the most important findings of this research is that the PBO or Actuarial Liability depends fundamentally on a linear basis of two fundamental variables in the quantification of Social Benefits -- the Guarantees and the Social Benefits to Pay (PSP).

The level of contribution to the determination of the PBO is the guarantee (what the company has already accumulated for the payment). This level of contribution is in the order of 39.42% and the part relative to PSP Social Benefits to be paid is 27.71% with a statistic of adjusted-R squared of 90.16%. It appears only two predictive variables that the model explains reasonably -- the corresponding actuarial liability for the differential of benefits.

Realistically, companies are often required to give quick estimates, and they find it quite difficult to quantify an order of this magnitude. This model succeeds in filling this informational gap at the 4 percent interest rate. Areas of future research include repeating the study with other real rates to see if there is structural stability of the model.

The potential of a structural robust model is a very attractive option in situations where managers are tasked to provide estimates in relatively short order, given that the two variables highlighted are generally readily available for most organizations.

References

- [1] World Economic Forum, "We'll Live to 100 – How Can We Afford It?"
- [2] Gillespie, Patrick, "Half the Venezuelan Economy has Disappeared." @CNNMoney. Accessed April 12, 2018. <http://money.cnn.com/2018/01/25/news/economy/venezuela-imf/index.html>.
- [3] Alex, G.O., "Government Pensions in Venezuela Used to Mean Security in Old Age — but not any more," *Miami Herald*, November 4, 2017.
- [4] Chen, H., MacMinn, R. D. and Sun, T. (2017), "Mortality Dependence and Longevity Bond and Insurance. Pricing: A Dynamic Factor Copula Mortality Model With the GAS Structure," *Journal of Risk and Insurance*, 84: 393-415. 2017.
- [5] Njenga, C. & Sherris, M., "Longevity Risk and the Econometric Analysis of Mortality Trends and Volatility." *Asia-Pacific Journal of Risk and Insurance*, 5:35-55. 2011.
- [6] Query, J.T. and E. Diz Cruz, "A Retroactive Impact Analysis of Changes in Social Benefits: The Case of Venezuela," *International Journal of Business and Social Research*, 5:87-104. 2015
- [7] Bisetti, E., Favero, G.N., and Tebaldi, C., "A Multivariate Model of Strategic Asset Allocation with Longevity Risk," *Journal of Financial and Quantitative Analysis*, 52: 2251-2275. 2017.
- [8] Brooks, S. (2007). Globalization and Pension Reform in Latin America. *Latin American Politics and Society*, 49, 31-62.