Ciencia y Tecnología, 30(2): 16-26, 2014
ISSN: 0378-0524

COMPARISON OF NELSON-SIEGEL AND SVENSSON MODELS IN THE OPTIMIZATION OF YIELD CURVES IN THE COSTA-RICAN MARKET

Eduardo Piza[1], Javier Trejos[1], y Esteban Bermúdez[2‡]*

[1]Centro de Investigación en Matemática Pura y Aplicada (CIMPA) y Escuela de Matemática, Universidad de Costa Rica, San José, Costa Rica.
[2]Escuela de Matemática, Universidad de Costa Rica, San José, Costa Rica.
[‡]Junta de Pensiones y Jubilaciones del Magisterio Nacional, San José, Costa Rica

Recibido 6 de diciembre del 2014; aceptado 31 de diciembre del 2014

Abstract

In a study of optimization of yield curves for zero-coupon and forward rates, we compare the Nelson-Siegel and Svensson models for a set of observed prices of bonds of Government of Costa Rica. In both models the problem consists of fitting the yield curve and selecting sub-optimal parameters of functions in a non-linear family. We obtain better results using the Nelson-Siegel model, despite the fact that the Svensson model is a generalization of the former. We discuss the method, some results obtained and an implementation of these methods in the Costa-Rican stock-exchange market.

Resumen

Una familia de curvas de rendimiento cero cupón se estudió a partir de un conjunto de precios observados en el mercado costarricense y se compararon los modelos de Nelson-Siegel y de Svensson. En ambos modelos el problema consiste en ajustar la curva de rendimientos seleccionando los parámetros suboptimales de una familia no lineal de funciones. Los mejores resultados se obtuvieron con el modelo de Nelson-Siegel, a pesar de que el de Svensson es una generalización del primero. Se discute la metodología empleada así como algunos resultados obtenidos y la implementación de estos métodos en el mercado de la Bolsa Nacional de Valores de Costa Rica.

Key words: Non-linear optimization, yield curves, Nelson-Siegel model, Svensson model.

Palabras clave: Optimización no lineal, curvas de rendimiento, modelo de Nelson-Siegel, modelo de Svensson.

I. INTRODUCTION

The stock market in Costa Rica is smaller than other stock markets in the world. Most transactions there come from Costa-Rican government bonds in colones or in USA dollars; transactions involving private stocks are infrequent. For mainly this reason, in Costa Rica it is important to have a zero-coupon curve.

The *Bolsa Nacional de Valores* (BNV, the National Stock Exchange in Costa Rica) makes a daily report of the prices of various stocks, bonds and shares of the Costa-Rican market. To make this report, BNV takes into account all negotiations of the day, applies varied filters and separates the stocks into categories [7], [8].

One characteristic of the Costa-Rican market is that it has few transactions per day, which creates a problem of price observation.

For year 2008, Costa Rica has two yield curves, one for the colon[1] and another for the dollar. Any yield curve in Costa Rica measures a rate of return; the observation points are the rates of

[*]Corresponding author: javier.trejos@ucr.ac.cr

return of government bonds in a list. If a bond is not negotiated on a given day, the point is an average of the preceding five sessions; if the bond has no price for these past five sessions, the point has the same value of the last day. For any bond that is excluded from this list, its value is based on linear interpolation of observation points.

In Section 2 we summarize the models studied, the criteria and the methods used; section 3 contains the experimental results., preceding conclusions in Section 4.

II. ESTIMATION OF YIELD CURVES

FUNCTIONAL MODELS

We seek to construct a zero-coupon yield curve for the Costa-Rican bond market[2] [16], [17], [18]. If instrument k pays n coupons at times $0 < t_1 < \cdots < t_n$ and principal p is paid at t_n, the theoretical value of the instrument, or its price, is

$$V_k = r \sum_{i=1}^{n-1} e^{-z(t_i)t_i} + (r+p)e^{-z(t_n)t_n}$$

The objective is to derive a curve $z(t)$ such that theoretical values $\hat{V}$ best approximate the real values in the market, through minimizing quality criterion $E^2(V, \hat{V})$. Among the most widely used models for the construction of a yield curve $z(t)$ are [14], [19] the following.

- The Nelson-Siegel model [20] has five parameters -- β_0, β_1, β_2, λ_1, λ_2 -- that belong to set Θ, and functional expression

$$z(t) = z_\Theta(t) = \beta_0 + \beta_1 \left(\frac{1 - e^{-\lambda_2 t}}{\lambda_2 t} \right) + \beta_2 \left(\frac{1 - e^{-\lambda_2 t}}{\lambda_2 t} - e^{-\lambda_2 t} \right).$$

- The Nelson-Siegel & Svensson model, known also as Svensson [22], has six parameters -- $\Theta = \beta_0, \beta_1, \beta_2, \lambda_1, \lambda_2$ -- and functional form

$$z(t) = z_\Theta(t) = \beta_0 + \beta_1 \left(\frac{1 - e^{-\lambda_1 t}}{\lambda_1 t} \right) + \beta_2 \left(\frac{1 - e^{-\lambda_1 t}}{\lambda_1 t} - e^{-\lambda_1 t} \right) + \beta_3 \left(\frac{1 - e^{-\lambda_2 t}}{\lambda_2 t} - e^{-\lambda_2 t} \right).$$

- The Vašiček model [26], [27], which describes a stochastic process of the spot rate and which is useful when only few observations exist, is expressed as

$$dr(t) = c[\theta - r(t)]dt + \sigma dB(t),$$

  in which θ is unconditional spot rate r at time t, σ is the volatility of the spot rate, $c > 0$ is a reversion parameter of the mean and $B(t)$ denotes a standard Brownian motion.
- Spline models [3] use a base of B-splines or cubic splines; for example, the Merrill-Lynch exponential splines (MLES) are widely used.
- The trinomial tree comes from Hull & White [13].

---

[1] *Colon* is the national currency in Costa Rica.

[2] The bond market is also known as the debt market or credit market

In the Costa-Rican stock and share market, there have been few studies of the zero-coupon yield curve. In [2], the authors studied the problem focused on the Svensson model, following suggestions [5] and [6]; they implemented (in Matlab) the Svensson model using predefined routines, but the problem of local minima of the model was not tackled. Those authors did not recommend the use of MLES in Costa Rica as data are insufficient; the model would hence be over-estimated. They did not analyze the possibility of using the Nelson-Siegel or $Vašíček$ models, nor trinomial trees.

According to [1], [6], [9], [11], [25], [21], Table 1 shows the use of models for zero-coupon yield curves in some countries.[3]

Table 1. Models used in some countries and institutions.

Nelson-Siegel	Svensson	Splines
Belgium	Germany	Colombia
Colombia	Belgium	USA
Spain	Canada	Japan
USA	Spain	Norway
Finland	USA	UK
France	France	World Bank
Italy	Norway	
UK	Sweden	
	Switzerland	
	UK	

OPTIMIZATION CRITERIA

Several optimization criteria might be used to evaluate the parameters in a functional yield curve. For the Nelson-Siegel and Svensson models, a criterion of least squares is written as

$$E^2(V,\hat{V}) = \sum_{k=1}^{K} w_k (V_k - \hat{V}_k)^2$$

in which w_k are weighting factors. Nelson-Siegel and Svensson models generally use weights based on the duration of the instrument, whereas in [2] the authors use weights defined as

$$w_k = \frac{1}{B_k(1+T(k))},$$

in which B_k is the bid-ask spread that eliminates the effect of variability in some instruments (it is the difference between the offer of average sales and the offer of average purchases), and $T(k)$ is the number of preceding days in which instrument k was transacted; this formula enables the use of historical data, for which it is hence easier to find titles in diverse terms.

The sum of these weights is not unity, as in the general case; to divide by the sum would suffice to obtain standard weights, but for optimization purposes this condition is expected not to affect the results.

Other criteria that might be used, depending on the model, include [2]

[3] In some countries multiple models have been used.

- maximum likelihood, with a well known solution in the Gaussian case, but which is unsuitable for our scarce data;
- L_1 criterion $E(V,\hat{V}) = \sum_{k=1}^{K} w_k |V_k - \hat{V}_k|$;
- Bayesian approach, which depends also on parameter estimation;
- neural nets, models in a non-parametric family.

OPTIMIZATION METHODS

As stated, the problem of evaluating parameters is a parametric problem in non-linear regression, namely minimizing weighted least squares as a criterion to evalutate parameters Θ in a set that define curve $z(t)$ such that $\hat{V} = f(z_\Theta(t))$; $z(t)$ is a non-linear function that cannot be made linear.

Among the most widely used methods to solve this problem, we cite these [10], [23]:

- *Gauss-Newton method*, which approximates $z_\Theta(t)$ by a first-order Taylor polynomial around a point (an initial estimate). This method makes a sequence of iterations based on multiple linear regressions until convergence is attained, but convergence is not assured.
- *method of gradient descent*, in which an iterative search of the direction of steepest descent proceeds on calculating a vector of derivatives, or gradient, step by step, until convergence, which is also not assured.
- *Marquardt method* [15], which combines the preceding two methods.

Apart from the possibility of a failure to converge, there is a problem of local minima. These methods are based on local search strategies; a functional form, depending on the model, might have multiple local minima. We illustrate the presence of local minima in the results of our experiments.

Variants of the Gauss-Newton method, such as the projected Newton method [4], can be used (in Matlab©), and enable avoiding problems with derivatives of the cost function; at the same time it accelerates the convergence.

The generalized reduced-gradient method (in Excel Solver©) has also advantages over the classical Gauss-Newton method.

The geometry of the parameter space must be taken into account. The Nelson-Siegel method entails only five parameters, the Svensson method six. In principle, a satisfactory solution is more difficult with the Svensson method than with the Nelson-Siegel one; we have proved this claim in the present case.

III. EXPERIMENTAL RESULTS

For experiments we used the colon bonds in Category 1 of 2006 March, shown in Table 2.

All instruments with coupon have periodicity 2; for a zero-coupon instrument we used periodicity 1 because its maturity period is less than, or equal to, one year. The data that have been used can be weighted according to the inverse of the duration; it is hence unnecessary to use any backward-average- indexed process in time, as in [2].

SOFTWARE

Hess [12] developed a program[4] (Microsoft Excel©) to compute zero-coupon yield curves with four models: Nelson-Siegel, Svensson, Vašíček and trinomial. Weights w_k, defined as the inverse of the durations, are real weights in the sense that their sum is equal to unity. We adapted this program to perform our experiments to compare the Nelson-Siegel and Svensson models in the Costa-Rican market. Hess built a dirty price using the bid-ask spread of the negotiations but this method is not used in Costa Rica. For experiments, because of the problem of price observation, an average dirty price was used, based on the volume and the seniority of the negotiation [8].

Table 2. Interest rates, date of maturity and price for bonds in colones.

Series	Interest Rate	Maturity	Price
data of 2006 March 1			
G290306	17.75%	03/29/2006	100.29%
BCCR111006	0.00%	10/11/2006	91.89%
BCCR111006	0.00%	10/11/2006	91.89%
N100107	0.00%	01/10/2007	88.83%
N100107	0.00%	01/10/2007	88.83%
D062003	17.90%	03/23/2007	103.30%
BCCR280307	18.00%	09/28/2006	103.30%
BCCR260907	18.00%	09/26/2007	104.65%
BCCR051207	16.02%	12/05/2007	101.17%
BCCR051207	16.02%	12/05/2007	101.15%
BCCR051207	16.02%	12/05/2007	101.15%
BCCR051207	16.02%	12/05/2007	101.15%
G260308	18.96%	03/26/2008	107.35%
BCCR30092009	18.75%	09/30/2009	107.75%
data of 2006 March 2			
BCCR260406	0.00%	04/26/2006	98.0020%
N120706	0.00%	07/12/2006	95.2987%
G270906	17.75%	09/27/2006	101.6233%
BCCR111006	0.00%	10/11/2006	91.9230%
BCCR280307	18.00%	03/28/2007	103.0700%
G260907	18.00%	09/26/2007	103.9600%
BCCR051207	16.02%	12/05/2007	101.1700%
G260308	18.96 %	03/26/2008	106.1160%
G300909	18.75 %	09/30/2010	107.6350%
G290910	16.75 %	09/29/2010	102.5800%
data of 2006 March 6			
BCCR120706	0.00%	07/12/2006	95.4310%
G270906	17.75%	09/27/2006	101.6467%
BCCR280307	18.00%	03/28/2007	103.0900%
BCCR30092009	18.75%	09/30/2009	107.6500%
G260912	17.30 %	09/26/2012	105.1100%

[4] The program is available at website *http://www.mngt.waikato.ac.nz/kurt/.*

RESULTS

The experiment was designed to evaluate the difference between the presence or absence of coupons in the instruments, and whether the spot rate is fixed or estimated in the model as a result of an optimization. Preliminary experiments showed that considering this difference and evaluating it might be of interest. There are hence two factors each with two levels, for a comparison of two models:

- factor 1, *Coupons* -- instruments with coupons and without coupons.
- factor 2, *Spot rate* -- fixed or variable, i.e., spot rate fixed or deterministic, and a rate that is a result of the optimization, estimated with the model.

The National Stocks and Shares of Costa Rica provided three real data sets for 2006 March 1, 2 and 6 (see Appendix). Each is considered a repetition of the experiment. For each data set and for each combination of factor levels, we executed, 100 times, each Excel model from randomly selected initial points, yielding in total $3 \times 2 \times 8 \times 100 = 2400$ executions of the program.

Results are presented in three tables, one for each data set: results for March 1 in Table 3, for March 2 in Table 4 and Table 4 for March 6. Each table is separated in two sections: the top contains the results for a fixed rate, the bottom for a variable rate. Models are separated in columns, Nelson-Siegel at left and Svensson at right; each contains also two columns, the first for instruments with coupons, the second for instruments without coupons. For each table we report

- the error of the best solution obtained in 100 executions, called the error at minimum;
- the number of times that the optimization procedure converged called the number of successes;
- the number of solutions of which the value of the error function is at a distance no greater than 1% from the value of error at the minimum; this number is called Percentile 1;
- the number of solutions of which the error function value is at a distance no greater than 5 % from the value of error at minimum; this number is called Percentile 5, and
- the duration of execution for 100 runs.

Table 3. Experimental results for data of March 1.

Fixed rate	Nelson-Siegel		Svensson	
	With coupon	Without coupon	With coupon	Without coupon
Error at minimum	0.0001310	0.0005668	0.0017126	0.0049300
Number of successes	95	88	92	94
Percentile 1	15	19	1	31
Percentile 5	58	50	1	35
Total duration	ND	ND	ND	ND
Variable rate	Nelson-Siegel		Svensson	
	With coupon	Without coupon	With coupon	Without coupon
Error at minimum	0.0001177	0.0004223	0.0010456	0.0040077
Number of successes	92	91	85	74
Percentile 1	34	17	1	1
Percentile 5	79	81	2	1
Total duration	ND	ND	ND	ND

From Tables 3, 4 and 5, we remark that:

- The value of the error function is generally less for the Nelson-Siegel model than for the Svensson model. If the model is executed several times and only the best solution is retained, generally the Nelson-Siegel model yields a superior fit.

- The number of successes is comparable for both methods; both methods might diverge a similar number of times. The difference for the number of successes between using a coupon or not, or between fixed or variable rate, is insignificant.

- The percentiles indicate the number of solutions near the best solution obtained. In some cases the Nelson-Siegel model was superior, in other cases the Svensson model. For any model, many solutions are remote from the best solution, which reflects the fact that the function finds typically finds local minima.

- The Nelson-Siegel model executes more rapidly than the Svensson model, but there is no difference of time between the use of coupons or not, or between fixed or variable rate.

Table 4. Experimental results for data of March 2.

Fixed rate	Nelson-Siegel		Svensson	
	With coupon	Without coupon	With coupon	Without coupon
Error at minimum	0.0000494	0.0000156	0.0077321	0.0018325
Number of successes	74	95	92	96
Percentile 1	2	5	3	1
Percentile 5	3	6	92	1
Total duration	4′8′′	3′55′′	5′20′′	5′13′′
Variable rate	**Nelson-Siegel**		**Svensson**	
	With coupon	Without coupon	With coupon	Without coupon
Error at minimum	0.0000516	0.0000257	0.0075450	0.0012706
Number of successes	88	88	89	93
Percentile 1	3	3	10	1
Percentile 5	12	3	65	5
Total duration	4′42′′	4′48′′	5′58′′	6′16′′

Table 5. Experimental results for data of March 6.

Fixed rate	Nelson-Siegel		Svensson	
	With coupon	Without coupon	With coupon	Without coupon
Error at minimum	0.0000374	0.0000020	0.0098729	0.0017099
Number of successes	71	73	83	81
Percentile 1	3	2	8	1
Percentile 5	5	2	82	1
Total duration	4′02′′	4′03′′	4′49′′	4′51′′
Variable rate	**Nelson-Siegel**		**Svensson**	
	With coupon	Without coupon	With coupon	Without coupon
Error at minimum	0.0000464	0.0000023	0.0081499	0.0025565
Number of successes	87	84	77	90
Percentile 1	2	10	3	1
Percentile 5	2	10	5	1
Total duration	4′02′′	4′01′′	6′23′′	4′58′′

IV. CONCLUSIONS AND FURTHER RESEARCH

In some cases the methods fail to converge; the methods and their results clearly depend on initial estimates: they might find local minima.

The Nelson-Siegel method is better adapted and easier to use than the Svensson method. The latter is supposed to be a generalization of the former, and is expected theoretically to find the same solutions, but in practice that result does not occur. The inclusion of additional parameters impedes finding better solutions, makes the work of estimation more difficult and is generally useless. The Nelson-Siegel method hence obtains better solutions than the Svensson method, and its superior fits require less time.

The problem of local minima that arises in both methods is perfectly appreciated according to Tables 2, 3 and 4. In a strategy of multiple initiations, several executions of the methods are hence recommended, as the least that one should make. A better strategy, however, would be to apply combinatorial optimization metaheuristics in non-linear regression to improve the existing methods. In this sense, the results obtained in [24] or [28] are promising and we shall explore this approach in the near future.

V. REFERENCES

1. Abad Romero, P.; Robles Fernández, M.D., Estructura temporal de los tipos de interés: teoría y evidencia empírica, *Revista Asturiana de Economía*, **2003**, 27, 7–47.
2. Barboza, L.; Ramírez, J.A.; Víquez, J.J., Curva de rendimientos cero cupón para el mercado costarricense, Preprint, Sección de Riesgos, Banco Nacional de Costa Rica, 2005, 25 pp.
3. Barboza, L., Modelo de Estimación de Pérdidas de Mercado en Swaps de Tasas de Interés, Tesis de maestría en Matemática Aplicada, , Universidad de Costa Rica: Ciudad Universitaria Rodrigo Facio, 2007.
4. Bertsekas, D.P.; Gafni, E.M., Projected Newton methods and optimization of multicommodity flows, *IEEE Trans Automatic Control*, **1983**, 28(12), 1090-1096.
5. Bolder, D.J.; Streliski, D., Yield curve modeling at the Bank of Canada, Technical Report No. 84, Bank of Canada, 1999, 69 pp.
6. Bolder, D.J.; Gusba, S., Exponentials, polynomials and Fourier series: more yield curve modeling at the Bank of Canada, Working Paper 2002–29, Bank of Canada, 2002, 81 pp.
7. Bolsa Nacional de Valores, Metodología de Cálculo Vector de Precios, 2006, 30 pp.
8. Bolsa Nacional de Valores, Estimación hecha por la BNV según el modelo de Nelson & Siegel (1987) para una curva de rendimiento cero cupón, 2007, 8 pp.
9. Debt Management Office, The DMO's yield curve model, United Kingdom Debt Management Office, July, 2000, 11 pp.
10. Draper, N.R.; Smith, H., *Applied regression analysis*. John Wiley & Sons: New York, 1968.
11. Filipovic, D., Exponential-polynomial families and the term structure of interest rates, *Bernoulli*, **2000**, 6(6), 1081–1107.
12. Hess, K., Bond relative value models and term structure of credit spreads: a practitioners approach, Working Paper, University of Waikato, Hamilton, New Zealand, 2004, 46 pp.
13. Hull, J.C., *Options, futures, and other derivatives*, 5th. Edition. Prentice-Hall: Englewood Cliffs NJ, 2002.
14. Ilmanen, A., Market rate expectations and forward rates, *The Journal of Fixed Income*, **1996**, 6(2), 8-22.
15. Marquardt, D.W., An algorithm for least squares estimation of nonlinear parameters, *Journal of the Society for Industrial and Applied Mathematics*, **1963**, 11(2), 431–441.
16. Matarrita, R., Determinación de la estructura temporal de tasas de interés en el mercado de valores costarricense: una propuesta, Documento de Trabajo 01–2003, Bolsa Nacional de Valores, San José, Costa Rica, 2003, 13 pp.
17. Matarrita, R., Definición de la curva soberana y la fragmentación del mercado, Documento de Trabajo 04–2003, Bolsa Nacional de Valores, San José, Costa Rica, 2003, 10 pp.
18. Matarrita, R., Apuntes para la construcción de una relación funcional rendimiento-plazo para el mercado de valores costarricense, Documento de Trabajo 09–2003, Bolsa Nacional de Valores, San José, Costa Rica, 2003, 13 pp.
19. Molinare, A., Estructura y dinámica de tasas de interés reales en Chile: información contenida en los pagarés reajustables con pagos en cupones del Banco Central, Tesis de maestría, Universidad Católica de Chile, Santiago, 2002.
20. Nelson, C.; Siegel, A., Parsimonious modelling of yield curves, *Journal of Business*, **1987**, 60(4), 473–489.
21. Place, J., Basic Bonds Analysis. Handbooks in Central Banking No. 20, Bank of England, 2000.

22. Svensson, L.E. Estimating and interpreting forward interest rates: Sweden 1992–94, Working Paper No. 114, International Monetary Fund, 1994.
23. Tomassone, R.; Audrain, S.; Lesquoy, E.; Millier, C. *La régression. Des nouveaux regards sur une ancienne méthode statistique*. Masson: Paris, 1992.
24. Trejos, J.; Villalobos, M., Optimización son sobrecalentamiento simulado en regresión no lineal: algoritmo y software, *Investigación Operacional*, **2000**, 21(3), 236–246.
25. Vásquez, D.M.; Melo, L.F., Estimación de la estructura a plazo de las tasas de interés en Colombia por medio del método de funciones B-spline cúbicas, Borradores de Economía, Banco de la República, Bogotá, 2002.
26. Vašiček, O.A., An equilibrium characterization of the term structure, *J. Finan. Econom*., **1977**, 5(2),177–188.
27. Vašiček, O.A.; Fong, H.G., Term structure modeling using exponential splines, *Journal of Finance*, **1982**, 37(2), 339–348.
28. Villalobos, M.; Trejos, J.; de los Cobos, S., Aplicación de la búsqueda tabú en regresión no lineal, *Revista de Matemática: Teoría y Aplicaciones*, **2006**, 13(1), 70–92.

APPENDIX

DATA

Series	Interest Rate	Maturity	Price
data of 2006 March 1			
G290306	17.75%	03/29/2006	100.29%
BCCR111006	0.00%	10/11/2006	91.89%
BCCR111006	0.00%	10/11/2006	91.89%
N100107	0.00%	01/10/2007	88.83%
N100107	0.00%	01/10/2007	88.83%
D062003	17.90%	03/23/2007	103.30%
BCCR280307	18.00%	09/28/2006	103.30%
BCCR260907	18.00%	09/26/2007	104.65%
BCCR051207	16.02%	12/05/2007	101.17%
BCCR051207	16.02%	12/05/2007	101.15%
BCCR051207	16.02%	12/05/2007	101.15%
BCCR051207	16.02%	12/05/2007	101.15%
G260308	18.96%	03/26/2008	107.35%
BCCR30092009	18.75%	09/30/2009	107.75%
data of 2006 March 2			
BCCR260406	0.00%	04/26/2006	98.0020%
N120706	0.00%	07/12/2006	95.2987%
G270906	17.75%	09/27/2006	101.6233%
BCCR111006	0.00%	10/11/2006	91.9230%
BCCR280307	18.00%	03/28/2007	103.0700%
G260907	18.00%	09/26/2007	103.9600%
BCCR051207	16.02%	12/05/2007	101.1700%
G260308	18.96 %	03/26/2008	106.1160%
G300909	18.75 %	09/30/2010	107.6350%
G290910	16.75 %	09/29/2010	102.5800%
data of 2006 March 6			
BCCR120706	0.00%	07/12/2006	95.4310%
G270906	17.75%	09/27/2006	101.6467%
BCCR280307	18.00%	03/28/2007	103.0900%
BCCR30092009	18.75%	09/30/2009	107.6500%
G260912	17.30 %	09/26/2012	105.1100%