Financial Engineering and the Japanese Markets **2**: 169–179, 1995.

A Constrained Least Square Approach to the Estimation of the Term Structure of Interest Rates

HIROSHI KONNO and TORU TAKASE
Department of Industrial Engineering and Management, Tokyo Institute of Technology, Tokyo, Japan

Abstract. This paper proposes a new practical method for estimating forward rate curves using bond prices available in the market. It is intended to improve the least square estimation method proposed by Carleton and Cooper by imposing additional constraints to guarantee the smoothness of the forward rate curves. The resulting problem is a nonconvex minimization problem, for which we will propose an efficient algorithm for calculating an approximate optimal solution. Computational experiments show that this method can efficiently generate smooth forward rate curves without increasing the residual errors in terms of least square fitting. Also, we will compare this result with an alternative and more efficient constrained least absolute deviation method.

Keywords. Carleton–Cooper's method, least square estimation, nonlinear programming, forward rate curves, bond prices.

1. Introduction

The purpose of this article is to propose a new practical method for estimating the term structure of interest rates (forward rate curve) which plays a crucial role in the evaluation of bond prices. Given the term structure, we can calculate the mispricing, i.e., the difference between the theoretical price and the market price of individual bonds. If the market price of a certain bond is significantly lower (higher) than the theoretical price, we would be able to obtain an excess return by buying (selling) this undervalued (overvalued) bond.

Also, the term structure can be used to calculate such indexes as effective yield, yield to maturity, duration and convexity of each bond. These indexes are of particular importance when we evaluate a bond portfolio. Further, the forward rate curve plays an essential role in the evaluation of derivatives by way of Heath–Jarrow–Morton model (1990). Many interesting bond portfolio optimization models have been developed (See Konno and Inori, 1989) in recent years and in these models it will be also necessary to make an efficient estimation of the term structure to construct a portfolio with smaller risk and larger return.

In 1976, Carleton and Cooper, in their pioneering paper proposed a least square method to estimate the term structure from the prices of bonds in the market. It turned out that the calculated forward rate tends to become unstable toward the end of the horizon when applied to the Japanese bond market. In fact, it fluctuates more than ten percent per year as we proceed far into the future. This is unrealistic from the practical point of view, since a large fluctuation of interest rate is usually associated with a significant change of the economic condition, which cannot be projected far into the future.

Thus this method should be replaced by a more stable method using a smooth polynomial function interpolating fewer data with smaller errors such as the ones calculated from government discount bonds (Chambers et al., 1984; McCulloch, 1971, 1975). Unfortunately, however, we have only a few discount government bonds in Japan, most of which are of short maturity, so that this interpolation method may not be reliable enough to be used for practical purposes. Thus we modified the Carleton–Cooper method by imposing appropriate restrictions on the fluctuation of forward rates so that the resulting forward rate curve is smooth enough.

In the next section, we will summarize the least square method developed by Carleton and Cooper. Section 3 will be devoted to several schemes to generate smoother forward rate curve and to a few additional modifications of the standard model. Numerical results of these methods will be presented in Section 4.

2. Classical Least Square Method

Let there be n types of bonds $B_j(j=1,\cdots,n)$ in the market. Also, let T_j be the maturity, c_j the coupon/period, f_j the face value and p_j the market price of the bond B_j. The theoretical price P_j of B_j is given by the formulat below.

$$P_j=c_j\left\{\frac{1}{(1+i_1)}+\frac{1}{(1+i_1)(1+i_2)}+\cdots+\frac{1}{(1+i_1)(1+i_2)\cdots(1+i_{T_j})}\right\}$$
$$+f_j\frac{1}{(1+i_1)(1+i_2)\cdots(1+i_{T_j})}, \tag{1}$$

where i_t is the forward rate during period t.

Let T be the time horizon of the estimation process under consideration and let

$$y_t=\frac{1}{(1+i_t)}, \quad t=1,\ldots,T. \tag{2}$$

Also, let

$$c_{jt}=\begin{cases} c_j, & t=1,\ldots,T_{j-1}, \\ c_j+f_j, & t=T_j, \\ 0, & t=T_{j+1},\ldots,T. \end{cases} \tag{3}$$

Then the equation (1) can be rewritten as follows

$$P_j=c_{j1}y_1+c_{j2}y_1y_2+\cdots+c_{jT}y_1\cdots y_T, \quad j=1,\ldots,n. \tag{4}$$

Further, let

$$z_t=y_1y_2\cdots y_t, \quad t=1,\ldots,T. \tag{5}$$

Then we have a linear expression

$$P_j = c_{j1}z_1 + c_{j2}z_2 + \cdots + c_{jT}z_T, \quad j = 1, \ldots, n \tag{6}$$

which is equivalent to (4). Thus if we obtain a good estimate of $z_t (t = 1, \ldots, T)$, then we can recover y_t by using the relation

$$y_t = \frac{z_t}{z_{t-1}}, \quad t = 1, \ldots, T, \tag{7}$$

where $z_0 = 1$. The forward rate i_t is calculated by the formula (2).

Since $i_t \geq 0$ for all t, variables y_t should be always less than or equal to one. Viewing (7), z_t's have to satisfy the relation

$$1 \geq z_1 \geq z_2 \geq \cdots \geq z_T > 0. \tag{8}$$

The difference between the theoretical price P_j and the market price p_j

$$e_j = \sum_{t=1}^{T} c_{jt}z_t - p_j, \tag{9}$$

is called the mispricing of the bond B_j.

Thus the standard least square estimation leads to the following quadratic programming problem:

$$\left| \begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} \left\{ \sum_{t=1}^{T} c_{jt}z_t - p_j \right\}^2 \\ \text{subject to} & 1 \geq z_1 \geq z_2 \geq \cdots \geq z_T > 0, \end{array} \right. \tag{10}$$

which can be solved very fast by standard algorithms (Luenberger, 1984). This is the method proposed by Carleton and Cooper in 1976.

As noted in the introduction, this model leads to unstable forward rate sequences when applied to the bond data of the Japanese market. This is partly due to the fact, that the investor's projection of interest rates tends to become imprecise as we proceed toward the future. Also, sometimes a considerable instability is observed in the earlier stages of the entire horizon, which is considered to be a significant drawback of this approach.

In the next section, we will propose several modifications to ensure a more stable forward rate sequence.

3. Improvements of Carleton–Cooper Method

3.1. IMPROVEMENT TO GUARANTEE SMOOTHNESS

The first and the simplest modification is to add an upper bound constraint on the forward rate, i.e.,

$$i_t \leq i_{\max}, \quad t = 1, \ldots, T, \tag{11}$$

where $i_{\max}$ is the maximal allowable value (10%/year, say) of the forward rate. This constraint is equivalent to

$$z_T \geq dz_{T-1} \geq d^2 z_{T-2} \geq \cdots \geq d^{T-1} z_1 \geq d^T, \tag{12}$$

where $d=(1+i_{\max})^{-1}$.

The second modification is to impose smoothness conditions on forward rates. The most straightforward method is to add the following constraint;

$$|i_t - i_{t+1}| \leq \varepsilon, \quad t=1, \ldots, T-1. \tag{13}$$

However, this leads to an intractable nonlinear constraint in terms of z variables, i.e.,

$$\left| \frac{z_{t-1}}{z_t} - \frac{z_t}{z_{t+1}} \right| \leq \varepsilon, \quad t=1, \ldots, T-1. \tag{14}$$

To avoid this difficulty, we impose a constraint of the form

$$\frac{1}{1+\varepsilon} \leq \frac{1+i_{t+1}}{1+i_t} \leq 1+\varepsilon, \quad t=1, \ldots, T-1. \tag{15}$$

Note that this condition can be rewritten as follows

$$-\varepsilon(1+i_{t+1}) \leq i_{t+1} - i_t \leq \varepsilon(1+i_t), \quad t=1, \ldots, T-1.$$

Hence, (15) is almost equivalent to (13) by noting that $0 \leq i_t \leq i_{\max}$, where $i_{\max}$ is usually less than 0.05 when we take 6 months as one period..

Using the relation (7), the constraint (15) can be represented in terms of z variables as

$$\frac{1}{1+\varepsilon} \leq \frac{z_t^2}{z_{t-1} z_{t+1}} \leq 1+\varepsilon, \quad t=1, \ldots, T-1. \tag{16}$$

Thus we obtain the problem:

$$\left| \begin{array}{ll} \text{minimize} & f(z) \equiv \sum_{j=1}^{n} \left\{ \sum_{t=1}^{T} c_{jt} z_t - p_j \right\}^2 \\ \text{subject to} & 1 \geq z_1 \geq \cdots \geq z_T, \\ & z_T \geq dz_{T-1} \geq \cdots \geq d^{T-1} z_1 \geq d^T, \\ & \frac{1}{1+\varepsilon} \leq \frac{z_t^2}{z_{t-1} z_{t+1}} \leq 1+\varepsilon, \quad t=1, \ldots, T-1, \\ & z_0 = 1. \end{array} \right. \tag{17}$$

This is a convex quadratic programming problem with additional nonconvex constraints for which there exists no efficient algorithm to calculate a global minimum.

In the sequel, we will propose a practical algorithm for obtaining a good locally optimal solution. Let us introduce an auxiliary variable

$$x_t = -\ln z_t, \quad t=1, \ldots, T, \tag{18}$$

and represent the problem (17) using x variables as follows:

$$
\begin{array}{ll}
\text{minimize} & g(x) \equiv \sum_{j=1}^{n} \left\{ \sum_{t=1}^{T} c_{jt} \exp(-x_t) - p_j \right\}^2 \\
\text{subject to} & 0 \le x_1 \le x_2 \le \cdots \le x_T, \\
& x_T \le a + x_{T-1} \le 2a + x_{T-2} \le \cdots \le (T-1)a + x_1 \le Ta, \\
& -\alpha \le x_{t-1} - 2x_t + x_{t+1} \le \alpha, \quad t = 1, \dots, T-1, \\
& x_0 = 0,
\end{array}
\tag{19}
$$

where $\alpha = \ln(1+\varepsilon)$ and $a = -\ln d$. This is a linearly constrained nonlinear least square problem whose objective function is nonconvex.

It is not easy in general to find a globally minimal solution of a nonconvex minimization problem. However, we can construct an efficient method for calculating a good locally optimal solution using the special structure of the problem (19).

The first step is to solve a convex quadratic programming problem:

$$
\begin{array}{ll}
\text{minimize} & f(x) \equiv \sum_{j=1}^{n} \left\{ \sum_{t=1}^{T} c_{jt} z_t - p_j \right\}^2 \\
\text{subject to} & 1 \ge z_1 \ge \cdots \ge z_T, \\
& z_T \ge d z_{T-1} \ge \cdots \ge d^{T-1} z_1 \ge d^T.
\end{array}
\tag{20}
$$

Let z_t^0 $(t = 1, \dots, T)$ be its optimal solution. If z_t^0's satisfy the constraint of the problem (17), we are done. Let z_t^k be the feasible solution of (17) obtained at the kth iteration. We will linearize the objective function $f(z)$ of (17) around $x_t^k \triangleq -\ln z_t^k$ and solve the convex quadratic programming problem:

$$
\begin{array}{ll}
\text{minimize} & g_h(x) \equiv \sum_{j=1}^{n} \left\{ \sum_{t=1}^{T} c_{jt} \exp(-x_t^k)(1 - (x_t - x_t^k)) - p_j \right\}^2 \\
\text{subject to} & 0 \le x_1 \le x_2 \le \cdots \le x_T, \\
& x_T \le a + x_{T-1} \le 2a + x_{T-2} \le \cdots \le (T-1)a + x_1 \le Ta, \\
& -\alpha \le x_{t-1} - 2x_t + x_{t+1} \le \alpha, \quad t = 1, \dots, T-1, \\
& x_0 = 0.
\end{array}
\tag{21}
$$

Let $x^{k+1} = (x_1^{k+1}, \dots, x_T^{k+1})$ be an optimal solution of this problem. If $g(x^{k+1}) < g(x^k)$, then we go to the next iteration. If, on the other hand $g(x^{k+1}) \ge g(x^k)$, then we stop calculation. Since the feasible region of (21) is compact, there exists an accumulation point, say x^*, of the sequence $\{x^k\}$. Therefore the calculation is stopped if

$$
g(x^k) - g(x^{k+1}) < \delta \quad \text{and} \quad |g(x^{k+1}) - g_h(x^{k+1})| < \gamma, \tag{22}
$$

is satisfied for small enough $\delta, \gamma > 0$.

Convergence of this approximation procedure may not be fast in general situation. However the approximation is rather accurate in our problem since x_t has to satisfy the condition

$$0 \le x_t \le T \ln(1+\varepsilon_{\max}) \le 1,$$

by noting that $T \le 20$ and $\varepsilon \le 0.05$ and usually much less in most circumstances.

It is remarked that if we choose ε small enough, we will have smoother forward rate curve. Instead, the calculated minimal value of the sum of the squares would increase as ε decreases. The problem is therefore essentially multi-objective and we need to have compromise between the smoothness of the forward rate curve and the sum of the squares of mispricings by choosing an adequate level of ε.

3.2. Further Improvements

Instead of minimizing the sum of the squares of mispricing, we may minimize the sum of the absolute value of mispricing, i.e.,

$$\left|\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} \left| \sum_{t=1}^{T} c_{jt} z_t - p_j \right| \\ \text{subject to} & 1 \ge z_1 \ge z_2 \ge \cdots \ge z_T, \\ & z_T \ge d z_{T-1} \ge \cdots \ge d^{T-1} z_1 \ge d^T, \\ & \dfrac{1}{1+\varepsilon} \le \dfrac{z_t^2}{z_{t-1} z_{t+1}} \le 1+\varepsilon, \quad t=1, \dots, T. \end{array}\right. \tag{23}$$

By using a linear approximation to the objective function around $z_t^k = \exp(-x_t^k)$, we have the following problem:

$$\left|\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} \left| \sum_{t=1}^{T} c_{jt} \exp(-x_t^k)(1-(x_t - x_t^k)) - p_j \right| \\ \text{subject to} & 0 \le x_1 \le x_2 \le \cdots \le x_T, \\ & x_T \le a + x_{T-1} \le 2a + x_{T-2} \le \cdots \le (T-1)a + x_1 \le Ta, \\ & -\alpha \le x_{t-1} - 2x_t + x_{t+1} \le \alpha, \quad t=1, \dots, T-1, \\ & x_0 = 0. \end{array}\right. \tag{24}$$

It is well known that this problem can be reduced to a linear programming problem:

$$\left|\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} u_j + \sum_{j=1}^{n} v_j \\ \text{subject to} & \sum_{t=1}^{T} c_{jt} \exp(-x_t^k)(1-(x_t - x_t^k)) + u_j - v_j = p_j, \quad j=1, \dots, n, \\ & 0 \le x_1 \le x_2 \le \cdots \le x_T, \\ & x_T \le a + x_{T-1} \le 2a + x_{T-2} \le \cdots \le (T-1)a + x_1 \le Ta, \\ & -\alpha \le x_{t-1} - 2x_t + x_{t+1} \le \alpha, \quad t=1, \cdots, T-1, \quad x_0 = 0, \\ & u_j \ge 0, \ v_j \ge 0, \ j=1, \dots, n, \end{array}\right. \tag{25}$$

for which we can apply an efficient parametric simplex algorithm when we vary the value of α, or equivalently the value of ε parametrically.

Note that both least square and least absolute deviation estimates are unbiased estimates of e_j's if they are independently and normally distributed with mean zero and variance θ^2. In case e_j's do not satisfy these conditions, then the least absolute deviation estimation is more robust and computationally more efficient.

The underlying principle of the formulation above is to treat each bond equally regardless of the amount of transaction. We may instead assign larger weights to those bonds which are heavily transacted as proposed by Takamori and Shimizu (1994) in their recent article. This scheme leads to the following problem:

$$\left|\begin{array}{ll} \text{minimize} & F(z) \equiv \sum_{j=1}^{n} b_j \left| \sum_{t=1}^{T} c_{jt} z_t - p_j \right| \\ \text{subject to} & 1 \geq z_1 \geq \cdots \geq z_T, \\ & z_T \geq d z_{T-1} \geq \cdots \geq d^{T-1} z_1 \geq d^T, \\ & \dfrac{1}{1+\varepsilon} \leq \dfrac{z_t^2}{z_{t-1} z_{t+1}} \leq 1+\varepsilon, \quad t = 1, \ldots, T-1, \end{array}\right. \tag{26}$$

where b_j is the amount of B_j transacted in the market. They argue that this formulation can be interpreted as the minimization of total arbitrage in the bond market. Also, similar idea can be applied to the least square model, in which case we have to solve the followng problem:

$$\left|\begin{array}{ll} \text{minimize} & F(z) \equiv \sum_{j=1}^{n} b_j \left\{ \sum_{t=1}^{T} c_{jt} z_t - p_j \right\}^2 \\ \text{subject to} & 1 \geq z_1 \geq \cdots \geq z_T, \\ & z_T \geq d z_{T-1} \geq \cdots \geq d^{T-1} z_1 \geq d^T, \\ & \dfrac{1}{1+\varepsilon} \leq \dfrac{z_t^2}{z_{t-1} z_{t+1}} \leq 1+\varepsilon, \quad t = 1, \ldots, T-1, \end{array}\right. \tag{27}$$

In the next section, we will provide the results of numerical comparison of the standard formulation and several improvements proposed in this section.

4. Numerical Tests and Its Analysis

We calculate the forward rate curves using the market data of the 48 government bonds whose initial maturity is ten years. We choose 6 months as the length of one period and calculate the forward rate for twenty periods using NEC PC9821/i486$^{\text{TM}}$DX(33MHz).

Figures 1 and 2 show the results of the least square scheme (17) and the results of the least of the least absolute scheme (23) for various values of smoothing parameter α using the market data of September, 1992.

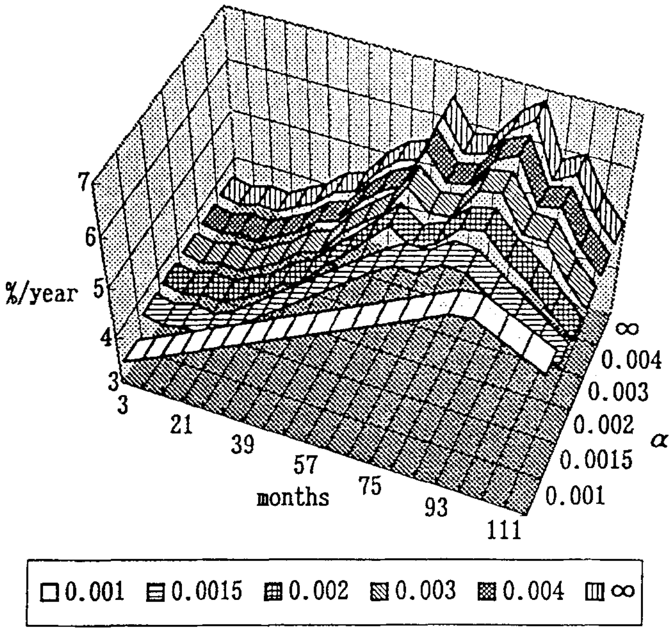

Fig. 1. The forward rate curve on September, 1992 with several α in QP.

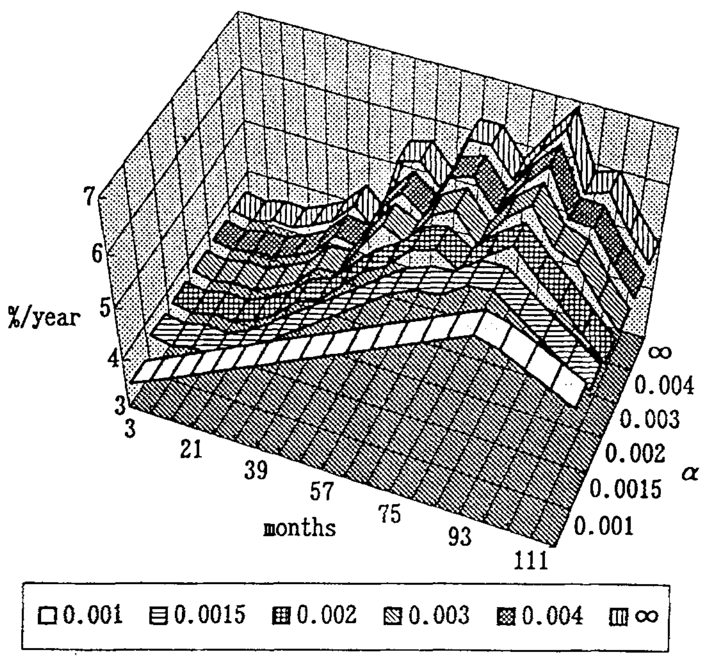

Fig. 2. The forward rate curve on September, 1992 with several α in LP.

It is seen from this that the forward rate curve becomes smoother as α decreases. We employed 10^{-5} as the value of convergence parameter δ throughout all experiments. Convergence condition was satisfied within 3 iterations. The CPU time until convergence never exceeded 60 seconds. The calculated solutions are very good locally optimal solutions, though there is no guarantee that they are globally optimal.

Figure 3 shows the forward rate curve for the market data of March, 1994, calculated by both the least square model (17) and the least absolute deviation model

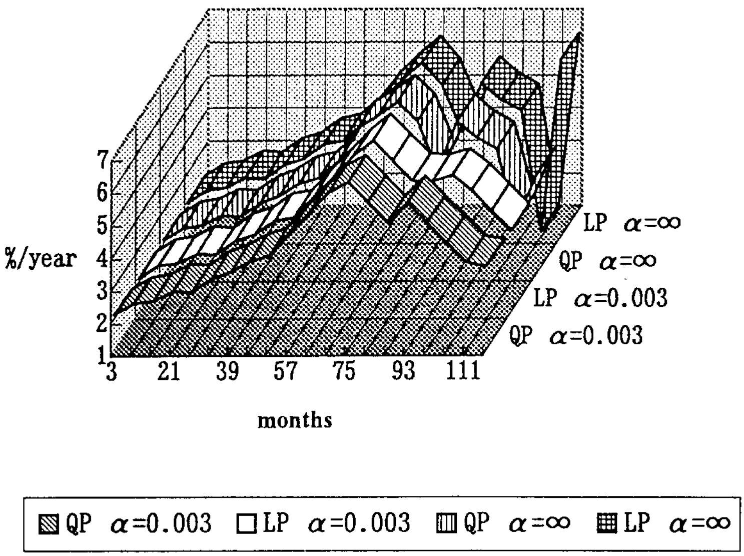

Fig. 3. The forward rate curve on March, 1994.

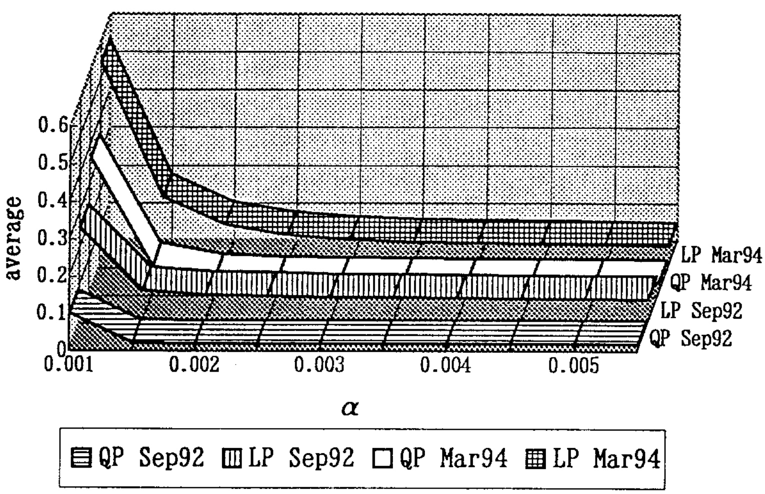

Fig. 4. The average of the squares of mispricing in QP or the average of the absolute deviations of mispricing in LP.

(23). We see from these curves that the least square scheme leads to somewhat smoother forward rate curves as expected.

The overall computation time was not significantly different. Least absolute deviation model (23) requires more iterations than (17), though it requires less computation per iteration. It appears that (23) is more efficient than (17) for the problem with larger T and/or n.

Figure 4 shows that average amount of mispricing e_j's for least square and least absolute deviation scheme for various values of α. It is observed from this that the amount of mispricing increases as α decreases, i.e., as we impose more strict restriction on the smoothness of the forward rate curve. Based upon a number of computations using market data, we conclude that $\alpha = 0.003$, or 0.3%/6 *months* fluctuation of forward rate, is an appropriate level of smoothing restriction, in which case the forward rate curve is sufficiently smooth while the amount of mispricing is almost the same as the case without restriction on smoothness, i.e., the case with $\alpha = \infty$. This conclusion applies to both least square and least absolute deviation schemes.

The calculated solution may not be a globally optimal solution of the problem (17). However, there is a good reason to believe that it is in fact globally optimal since we reached the same solution regardless of the choice of the starting feasible solution. (This point has to be investigated further.) One significant difference between (17) and (23) in regard to the pattern of mispricing is that more than 40% of the bonds are free from mispricing in the least absolute deviation scheme when α is ∞ while the mispricing is scattered among all bonds in the least square scheme.

Finally, Figures 5 and 6 show the comparison among the models (17), (23) and

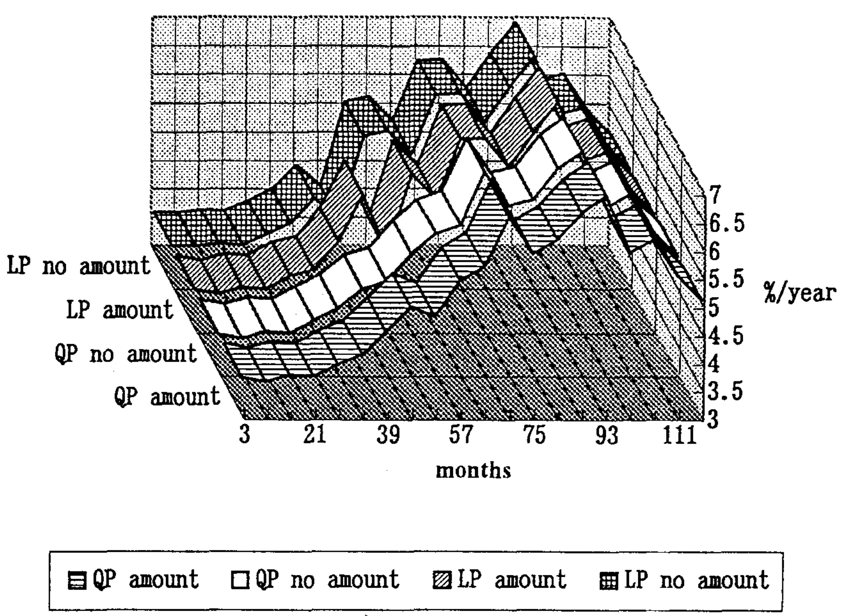

Fig. 5. The forward rate curve on September, 1992 ($\alpha = \infty$).

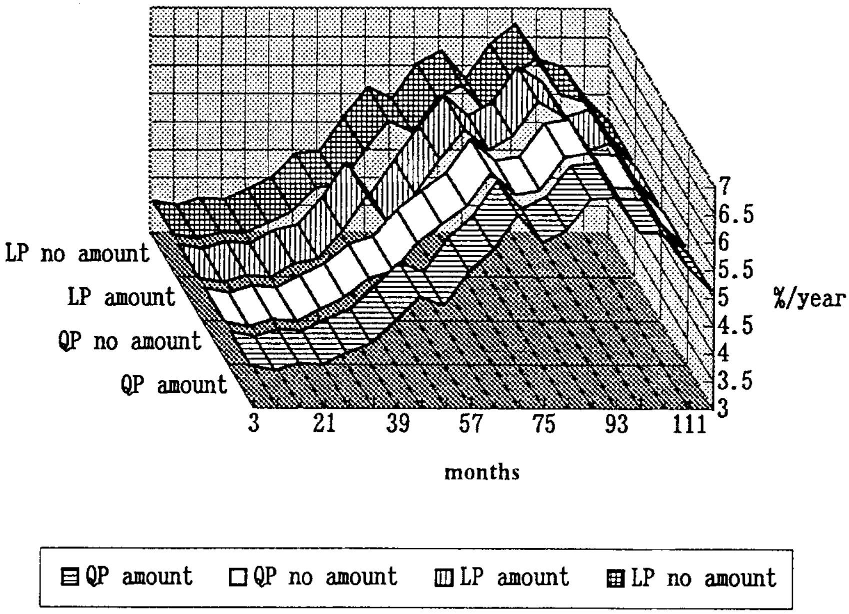

Fig. 6. The forward rate curve on September, 1992 ($\alpha = 0.003$).

(26), (27). There is certainly some difference, but it is hard to derive a definite conclusion from these figures, whether it is better to incorporate the amount of transaction into the objective function.

References

Carleton, W.T. and Cooper, I.A. (1976) 'Estimation and Uses of Term Structure of the Interest Rates', *Journal of Finance*, **31**, 1067 – 1083.

Chambers, D.R., Carleton, W.T. and Waldman, D.W. (1984) 'A New Approach to Estimation of the Term Structure of Interest Rates', *Journal of Financial and Quantitative Analysis*, **19**, 233–252.

Chvátal, V. (1983) *Linear Programming*, Freeman and Co.

Delbaen, F. and Lorimier S (1992) 'Estimation of the yield curve and the forward rate curve starting from a finite number of observations', *Insurance: Mathematics and Economics*, **11**, 259–269.

Elton, E.J. and Gruber, M.J. (1991) *Modern Portfolio Theory and Investment Analysis* (4th ed.), John Wiley & Sons.

Gill, P.E., Murray, W. and Wright, M.H. (1981) *Practical Optimization*, Academic Press.

Heath, D., Jarrow, R. and Morton, A. (1990) 'Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation', *Journal of Financial and Quantitative Analysis*, **25**, 419–440.

Konno, H. and Inori, M. (1989) 'Bond Portfolio Optimization by Bilinear Fractional Programming', *Journal of the Operations Research Society of Japan*, **32**, 143–158.

Luenberger, D.G. (1984) Introduction to Linear and Nonlinear Programming (2nd ed.), Addison-Wesley.

McCulloch, J.H. (1971) 'Measuring the Term Structure of Interest Rates', *Journal of Business*, 19–31.

McCulloch, J.H. (1975) 'The Tax-Adjusted Yield Curve', *Journal of Finance*, **30**, 811–830.

Sato, M. (1990) 'A Research Note on Estimation of the Term Structure of Interest Rates' (in Japanese), *MTEC Journal*, **3**, 85–116.

Takamori, Y. and Shimizu, Y. (1994) 'On the Efficiency of the Government Bond Market and the Estimation of the Term Structure of the Interest Rates' (in Japanese), *Proceedings of the JAFEE Summer Meeting*, 142–154.