UNIVERSITY OF WAIKATO

Hamilton
New Zealand

Modelling the Yield Curve with

Orthonormalised Laguerre Polynomials:

A Consistent Cross-Sectional and Inter-Temporal Approach

Leo Krippner

Department of Economics

Working Paper in Economics 2/03

September 2003

Leo Krippner
Reserve Bank of New Zealand
P.O. Box 2498
Wellington, New Zealand

Tel: +64 (04) 471-3605
Fax: +64 (07) 471-3800
Email: krippnerl@rbnz.govt.nz
http://www.rbnz.govt.nz

Abstract

This article proposes the orthonormalised Laguerre polynomial (OLP) model of the yield curve, a generic linear model that is both cross-sectionally consistent (that is, it reliably fits the yield curve at a given point in time), and inter-temporally consistent (that is, the cross-sectional parameters are shown to be consistent over time within the expectations hypothesis framework). The OLP model generalises the exponential-polynomial model for a single yield curve, as originally proposed by Nelson and Siegel (1987), and also allows for the simultaneous modelling of other same-currency yield curves that have instrument-specific differences (such as default risk), as in Houweling, Hoek and Kleibergen (2001). New Zealand data is used to illustrate the empirical application of the OLP model.

Keywords

JEL Classification

Acknowledgements

1 Introduction

Numerous models of the yield curve have evolved over past decades, essentially paralleling the increasing sophistication and variety of applications required by practitioners in modern financial markets. A popular group is that of parsimonious parametric cross-sectional models of the yield curve. These are applied to yield curve data observed at a point in time primarily to gauge a smooth and continuous function of yields by maturity, which may then be used for pricing purposes and/or gauging market interest rate expectations.

The literature on cross-sectional models is large. Broad classes (and recent examples) include ad-hoc functional forms (Echols and Elliot 1976), spline-based approaches (Lin 2002), orthogonal component approaches (Pham 1998), approaches based on equilibrium models of the yield curve (Hördahl 2000), and exponential-polynomial functional forms (i.e an exponential decay by maturity multiplied by a polynomial in maturity). The exponential-polynomial approach is originally proposed in Nelson and Siegel (1987), and is extended and revisited in Svensson (1994), Hunt (1995), Bliss (1997), Mansi and Phillips (2001), and Diebold and Li (2002). Models of this type are widely used by yield curve practitioners,[1] and perform very favourably in comparison with other yield curve models.[2]

Notwithstanding the large body of literature, a key issue that has long

---

[1] Bank for International Settlements (1999) notes that ten central banks (of twelve surveyed) routinely use either the Nelson and Siegel (1987) and/or the Svensson (1994) model as their primary method for analysing the yield curve. Kacala (1993), Barrett, Gosnell and Heuson (1995), Schich (1997), and Söderlind and Svensson (1997) are further examples of the practical application of exponential-polynomial models.

[2] See, for example, Dahlquist and Svensson (1994), Seppala and Viertio (1996), Bliss (1997), and Fergusson and Raymar (1998).

been over-looked in cross-sectional models is the explicit investigation of inter-temporal consistency. That is, given that the yield curve should embody expectations about the evolution of interest rates, the cross-sectional parameters used to describe the yield curve at each point in time should also be consistent with the expected evolution of those parameters over time. Secondly, with the notable exception of Houweling, Hoek and Kleibergen (2001), cross-sectional models have concentrated on modelling single yield curves, rather than simultaneously exploiting the additional information contained in same-currency yield curves that differ only by instrument-specific qualities, such as default-risk. Thirdly, proposed cross-sectional models tend to be prescriptive, rather than a generic class of model that may be extended as required.

This article proposes the orthonormalised Laguerre polynomial (OLP) model of the yield curve, which specifically addresses each of the issues noted in the previous paragraph. The article proceeds as follows: section 2 formalises the original Nelson and Siegel (1987) approach within the systematic class of orthonormalised Laguerre polynomials.[3] This allows the number of factors in a model for the base yield curve to be extended arbitrarily, as per the cross-sectional fit required by the user. Section 2.4 proposes a spread function for the OLP model that allows for the simultaneous estimation of other same-currency yield curves that are related to the base yield curve by a "spread function". Section 3 illustrates the practical application of the OLP model in its cross-sectional sense to New Zealand yield curve data, and also calibrates several

---

[3]Nelson and Siegel (1987) notes the potential for generalising to higher-order models using Laguerre functions. However, to the knowledge of the author, this has not yet been investigated in the literature.

parameters to facilitate the routine ongoing application of the model to New Zealand market data. Section 4 investigates the inter-temporal consistency of the OLP model specified in section 3 using the expectations hypothesis of the yield curve. This analysis predicts a convenient and sensible time-series representation for the OLP parameters. Section 5 reports the results of empirical tests against the theoretical predictions of the time-series model, again using New Zealand data. The conclusion suggests potential applications for the OLP model, and identifies areas for further related work.

2 The orthonormalised Laguerre polynomial model of the yield curve

This section firstly summarises orthonormalised Laguerre polynomials as a generic class of functions, and then proceeds to use them as the basis for the proposed generic model of the yield curve, including an allowance for spread functions. The final sub-section discusses the practicalities of estimating the model parameters.

2.1 Laguerre and orthonormalised Laguerre polynomials

As noted in standard texts,[4] Laguerre polynomials are of the form:

$$L_n(x) = \sum_{k=0}^{n} \frac{(-1)^k n! x^k}{(k!)^2 (n-k)!} \tag{1}$$

---

[4]See, for example, Courant and Hilbert (1953) pages 93 to 97, and Rainville and Bedient (1981) pages 395 to 396.

where n and k are integers.

Laguerre polynomials do not by themselves form an orthonormal set, but the related set of functions $\varphi_n(x) = \exp\left(-x/2\right) L_n(x)$ are orthonormal for the interval $0 \leq x < \infty$. Hereafter, expressions of the form $\varphi_n(x)$ are called orthonormalised Laguerre polynomials (OLPs), and the first three OLPs are $\varphi_0(x) = \exp\left(-x/2\right) \cdot 1$, $\varphi_1(x) = \exp\left(-x/2\right) \cdot (-x + 1)$, and $\varphi_2(x) = \exp\left(-x/2\right) \cdot \left(\frac{1}{2}x^2 - 2x + 1\right)$.

2.2 The OLP model for the forward rate curve

The generic OLP model for the forward rate curve is simply a linear combination of a constant and the specified number of OLPs, with some convenient reparameterisation. Specifically:

$$f(t, m) = \sum_{n=1}^{N} \beta_n\left(t\right) \cdot g_n(\phi, m) \tag{2}$$

where:

- $f(t, m)$ is the instantaneous forward rate curve as a function of maturity m (in years), observed at time t;

- N is the number of linear parameters in the OLP model;

- $\beta_n\left(t\right)$ are the linear parameters, at time t, associated with the functions of maturity $g_n(\phi, m)$, where n denotes each individual parameter and its associated function;

- $g_1(\phi, m) = 1$, and $g_n(\phi, m) = -\varphi_{n-2}(2\phi m)$ for $n > 1$; and

- ϕ is a fixed positive constant (which is responsible for the "natural shape" of the yield curve components for $n > 1$, since it alters the rate of decay of the exponential term in $\varphi_{n-2}(2\phi m)$).

The functions $g_n(\phi, m)$ may be referred to as forward rate "modes",[5] and convenient descriptive names for the first four modes are Level, Slope, Bow, and Wave. These modes are specified in equations 3a to 3d, and are illustrated in figure 1.

$$g_1(\phi, m) = 1 \tag{3a}$$

$$g_2(\phi, m) = -\exp(-\phi m) \tag{3b}$$

$$g_3(\phi, m) = -\exp(-\phi m)(-2\phi m + 1) \tag{3c}$$

$$g_4(\phi, m) = -\exp(-\phi m)(2[\phi m]^2 - 4\phi m + 1) \tag{3d}$$

[Figure 1 here]

An equivalent and often more convenient expression of the generic OLP model is the vector form:

$$f(t, m) = [\boldsymbol{\beta}_N(t)]' \, \mathbf{g}(\phi, m) \tag{4}$$

where:

---

[5] The term "mode" is analogous to its use in physics, where it refers to an integer-related series of solutions to a second-order differential equation. For example, the modes of a tensioned string (a one-dimensional wave equation) are individual musical notes and their harmonics. OLPs are solutions of the second-order differential equation; $x\frac{d^2\varphi_n(x)}{dx^2} + \frac{d\varphi_n(x)}{dx} + (n + \frac{1}{2} - \frac{x}{4})\varphi_n(x) = 0$. See Courant and Hilbert (1953) pages 328 to 331.

- $\boldsymbol{\beta}_N(t)$ is the N-vector containing the N coefficients $\beta_n(t)$;

- $\mathbf{g}(\phi, m)$ is the N-vector function containing the N modes $g_n(\phi, m)$.

2.3 The OLP model for the interest rate curve

The generic OLP model for the interest rate curve may be expressed as a linear combination of interest rate modes, i.e:

$$
\begin{aligned}
R(t, m) &= \sum_{n=1}^{N} \beta_n(t) \cdot s_n(\phi, m) & \text{(5a)} \\
&= [\boldsymbol{\beta}_N(t)]' \, \mathbf{s}(\phi, m) & \text{(5b)}
\end{aligned}
$$

where:

- $R(t, m)$ is the continuously compounding interest rate for maturity m, observed at time t;

- $s_n(\phi, m)$ are the interest rate modes associated with their respective forward rate modes; and

- $\mathbf{s}(\phi, m)$ is the N-vector function containing the N modes $s_n(\phi, m)$.

The interest rate modes are obtained from the corresponding forward rate modes in the usual manner, i.e $s_n(\phi, m) = \frac{1}{m} \int_0^m g_n(\phi, m) dm$. The first four interest rate modes are specified in equation 6 and are illustrated in figure 2:[6]

---

[6] The application of L'Hôpital's rule shows that equations 6b to 6d are properly defined at $m = 0$, with values of -1. They are also properly defined in the limit of infinite maturity, with values of zero.

$$s_1(\phi, m) \;\; = \;\; 1 \tag{6a}$$

$$s_2(\phi, m) \;\; = \;\; \frac{1}{\phi m}\left[\exp(-\phi m) - 1\right] \tag{6b}$$

$$s_3(\phi, m) \;\; = \;\; -\frac{1}{\phi m}\left[2\phi m \exp(-\phi m) + \exp(-\phi m) - 1\right] \tag{6c}$$

$$s_4(\phi, m) \;\; = \;\; \frac{1}{\phi m}\left[2\left[\phi m\right]^2 \exp(-\phi m) + \exp(-\phi m) - 1\right] \tag{6d}$$

[Figure 2 here]

2.4 Allowing for yield spreads in the OLP model

The discussion so far has implicitly assumed that, for a given currency, the yield on any interest rate instrument is only a function of the maturity of that instrument. However, this will not generally be the case if some instruments differ in other respects that influence their market value and hence their market yield. The most obvious example of such a difference is the default risk of the instrument issuer, but other factors such as relative instrument liquidity, coupon/tax effects, and market structure (such as "on-the-run" and "off-the-run" Treasury securities in the United States) may also cause a material yield difference even for interest rate instruments with identical maturities.

When modelling groups of interest rate instruments that have different intrinsic properties, Houweling, Hoek and Kleibergen (2001) advocates the joint estimation of the base yield curve with spread functions to allow for yield curves that sit above the base yield curve.[7] As noted in Houweling et al. (2001), the

---

[7] Those authors model the base yield curve as a cubic spline, and the spread function as a

advantage of this approach is that "appropriate" restrictions may be applied

to the spread functions; i.e the spread functions should equal zero at $m = 0$,

they should be "smooth", and they should be approximately monotonically

increasing, since this accords with the stylised facts, empirical evidence, and

theoretical models.[8]

The spread function proposed here for the OLP model is very simple:

$1 - \varphi_0(2\phi m)$ satisfies all of the restrictions noted in the previous paragraph.

Hence, the forward rate spread mode may be expressed as:

$$
g_{N+l}(\phi, m) = \begin{cases} 0 & \text{if cashflow is not from spread group } l \\ \\ 1 - \exp(-\phi m) & \text{if cashflow is from spread group } l \end{cases} \tag{7}
$$

where l ranges from 0 to L, with L being the number of different yield curves

related to the base yield curve by a spread function (so $l = 0$ represents the

base yield curve). Figure 3 illustrates the forward rate spread mode, and the

associated interest rate spread mode, which is:

$$
s_{N+l}(\phi, m) = \begin{cases} 0 & \text{if c/flow is not from s/g } l \\ \\ 1 + \frac{1}{\phi m}\left[\exp(-\phi m) - 1\right] & \text{if c/flow is from s/group } l \end{cases} \tag{8}
$$

[Figure 3 here]

In this sense, the spread function may be regarded as a "dummy variable",

---

quadratic spline.

[8]The instruments must be of a relatively high credit rating for these properties to apply.
See, for example, Jarrow, Lando and Turnbull (1997), and Helwege and Turner (1999).

and efficiency should be obtained in empirical applications due to the simultaneous estimation of the base yield curve parameters that are common to all yield curves of the same currency.

2.5 The $\textbf{OLP}(N, L)$ model

Adopting the notation that the number of modes in the base yield curve is N, and that an additional L spread functions are denoted using an index added to N, all forward rates and interest rates are respectively defined by the $\mathrm{OLP}(N, L)$ model as:

$$f(t, m) = \left[\boldsymbol{\beta}_{N+L}\left(t\right)\right]' \mathbf{g}(\phi, m) \tag{9}$$

$$R(t, m) = \left[\boldsymbol{\beta}_{N+L}\left(t\right)\right]' \mathbf{s}(\phi, m) \tag{10}$$

where the $(N + L)$-vectors $\mathbf{g}(\phi, m)$ and $\mathbf{s}(\phi, m)$ now contain N modes and L spread functions.

The OLP model for the base yield curve may be extended to an arbitrary number of modes, as required by the user. Forward rate modes for $n > 4$ are readily obtainable by following the specifications in sections 2.1, and the corresponding interest rate modes are readily obtainable from the forward rate modes, as in section 2.3, using software packages that allow for symbolic computation.[9] Similarly, the spread mode in section 2.4 may be applied to an

---

[9]For example, the empirical analysis in this article also uses the $N = 5$ "Ripple" mode. There does not appear to be a simple analytical or summation expression for the interest rate modes by n. However, the existence of analytical expressions for the interest rate modes

arbitrary number of different yield curves related to the base yield curve, as with the general specification in Houweling et al. (2001).

Note that the exponential-polynomial models of Nelson and Siegel (1987) and Diebold and Li (2002) are mathematically equivalent to the $\text{OLP}(3,0)$ model, and the specification by Hunt (1995) is equivalent to the $\text{OLP}(2,0)$ model. The models of Svensson (1994), Bliss (1997), Mansi and Phillips (2001) do not have an exact $\text{OLP}(N,L)$ analogue, since those models include exponential terms with two different rates of decay.[10]

2.6 Numerical estimation of the OLP model

The numerical estimation of the OLP model may be undertaken by minimising the squared residuals of the net present value of the fixed interest instruments represented by the yield curve data, i.e:

$$\text{Minimise: } \sum_{k=1}^{K} w_k \cdot \left\{ \varepsilon_k \left[\boldsymbol{\beta}_{N+L}\left(t\right), \phi \right] \right\}^2 \tag{11a}$$

$$\text{subject to: } \quad \varepsilon_k\left[\cdot\right] = \sum_{j=1}^{J[k]} a_{jk} \exp\left(-m_{jk} \cdot \left[\boldsymbol{\beta}_{N+L}\left(t\right)\right]' \mathbf{s}(\phi, m_{jk})\right) \tag{11b}$$

where:

---

is guaranteed using integration by parts, since the polynomial factors of the forward rate modes will terminate at zero under repeated differentiation, and the exponential factor is well-behaved under repeated integration.

[10] The OLP approach here could easily be extended to include different rates of exponential decay, but this would have direct implications for the analysis on inter-temporal consistency in sections 4 and 5. In any case, empirical cross-sectional comparisons of the Nelson and Siegel (1987) model against the alternatives show a questionable benefit for the addition of extra parameters. See, for example, Svensson (1994), Schich (1997), and Bank for International Settlements (1999).

- K is the number of fixed interest instruments used to define the yield curve/s;

- w_k is a weighting factor (which may be set to 1 if no weighting is required);

- $\varepsilon_k \left[\boldsymbol{\beta}_{N+L} \left(t \right), \phi \right]$ is a defined function of $\boldsymbol{\beta}_{N+L} \left(t \right)$ and ϕ, given the cash-flows defined in equation 11b;

- $J \left[k \right]$ is the number of fixed cashflows for instrument k;

- a_{jk} is the magnitude of the cashflow j of instrument k (defined to be negative for the settlement price, and positive for all cashflows beyond settlement); and

- m_{jk} is the maturity of the cashflow j of instrument k.

As noted in Söderlind and Svensson (1997), this is a standard approach to empirically estimating yield curve models using data for coupon-bearing instruments, but one subtle difference in the above formulation is to include the settlement price as one of the defined cashflows. This simplifies the functional form for the minimisation, and also retains market-standard settlement conventions rather than implicitly assuming that settlement occurs on the date that the yield curve data is observed.[11]

The nature of the functions used in the OLP model means that the estimation process may be conveniently undertaken using the Newton-Raphson method to optimise $\boldsymbol{\beta}_{N+L} \left(t \right)$ for a given ϕ. If ϕ is also to be freely estimated

---

[11] For example, market-quoted yields for New Zealand bonds assume settlement in two working days. The settlement price is defined by applying a market-standard bond pricing formula to the market-quoted yield.

for each cross-section of data (as with the intermediate step in section 3 to calibrate the value of ϕ), then the estimation may proceed in two-steps; a bisection search method for ϕ in conjunction with the Newton-Raphson method to find $\boldsymbol{\beta}$ for the given ϕ at each step of the bisection.[12]

Note that empirical work using exponential-polynomial models sometimes reports convergence problems in empirical estimation.[13] This was not evident in any of the applications that follow, and each cross-sectional estimation took less than one second on an office-standard computer, which suggests that the orthonormality of the OLP modes contributes to empirical robustness and efficiency.

3 The empirical application of the OLP model to New Zealand data

Having proposed the generic form of the OLP model, this section illustrates the empirical application to New Zealand government-risk and bank-risk yield data. Using the terminology of section 2.5, the appropriate model is the $\text{OLP}(N, 1)$, where the government-risk yield curve data is represented by N modes, and the bank-risk yield curve data is modelled simultaneously as a single spread curve relative to the government-risk yield curve (so L is fixed at 1).

Apart from providing a practical example of the general application of the OLP model, the aim is to calibrate values for N and ϕ while retaining an

---

[12]The articles for Canada and France in Bank for International Settlements (1999) describe similar two-step processes.

[13]See, for example, Hunt (1995), and Bank for International Settlements (1999).

adequate representation of each observation of yield curve data over the sample period. Fixing these parameters creates a linear model suitable for ongoing routine practical use, and a feasible example for the later work on inter-temporal consistency. The calibration of ϕ also has prior empirical support in Nelson and Siegel (1987), Barrett et al. (1995), and Diebold and Li (2002) (for United States data, with values of 7.30, 0.33, and 1.37, respectively).

3.1 Description of the data

The data are daily end-of-day market-quoted mid-rates for government-risk and bank-risk instruments for the period 10 July 1997 to 21 February 2002, as obtained from the Reserve Bank of New Zealand. The overnight cash-rate / Official Cash Rate (OCR) data are included in both the government-risk and bank-risk yield curve data groups, since it proxies the 1-day rate in both cases.[14]

The government-risk instruments are all liquid nominal government bonds that were on issue at the start of the data period, and those issued during the data period. Government bond yields with less than one year to maturity were excluded automatically in all cases for the reason that these bonds have the potential to, and often do, become "squeezed" (i.e "cornered" by several market participants) when they fall into the typical money-market maturity range.[15] The April 2003 bond was also excluded from 14 November 2001, since it was apparent at the time that it had became significantly squeezed, which

[14]An indicative interbank overnight cash-rate was compiled by the Reserve Bank up to 16 March 1999, and the OCR has been set by the Reserve Bank since 17 March 1999.

[15]Specifically allowing for the costs of repurchase transactions on these bonds may in principle avoid any "distortions" due to squeezes. However, reliable repurchase quotes are often not available for the New Zealand market, so exclusion is the only practical remedy.

soon led to an effective absence of market price-making in that bond.

The bank-risk instruments include 1 to 6-month bank-bill yields for the entire period, 1 to 7-year semi-annual swaps yields up to 26 September 1997, and to 10-years from 29 September (the date when a broker quote was first available for that maturity).

In summary, the data covers 1164 trading days; each trading day has data for a minimum of 20 instruments and a maximum of 22 instruments; the minimum number of government-risk instruments is seven; and the minimum number of bank-risk instruments is 14.

3.2   Applying the OLP model to the data

Figure 4 illustrates an example of the OLP$(3,1)$ yield curve, with $\phi = 1$, fitted to a cross-section of actual data. This example is chosen since the sharply opposing values of the Slope and Bow parameters nicely illustrate their intuitiveness; i.e the yield curve at this point in time may be described as "upwardly sloped" and "downwardly bowed". Figure 5 illustrates the associated yield and price residuals, which are used to gauge the goodness of fit for that cross-section. The maximum absolute price residual for this cross section is \$5,519, which is associated with the 15 July 2013 government bond. Note that since the estimation of the OLP model is based on price residuals, the yield residuals can become quite large for instruments with relatively small maturities. If a better fit to shorter-maturity yields is required (for example, a 1-day rate identical to the policy rate) then the weight applied to shorter-maturity yields may simply be increased to satisfy the users requirements. If yield residuals are to be

minimised, this may be achieved approximately by weighting the price residual of each instrument by the inverse of the basis point value for each instrument.

[Figure 4 here]

[Figure 5 here]

The calibration of N and ϕ is based on a non-parametric assessment of the absolute price residuals for each cross section across the entire sample period. This is because preliminary analysis indicated that price residuals for each cross section had a variance that fluctuated significantly over time (indicating heteroscedasticity) and also often displayed non-normality (according to Jarque-Bera tests), and so routine statistical techniques to determine appropriate restrictions on N and ϕ (e.g the F-test) would not be valid.[16]

Regarding the appropriate value for ϕ, figure 6 shows the results for $N = 4$ when ϕ is freely estimated for each cross section, denoted as ϕ_t. Before proceeding, note that the estimate of the Level parameter occasionally becomes practically "unreasonable" (compared to the 10-year swaps yield, a proxy for the long-maturity level of the yield curve) when ϕ_t falls below about 0.3, which offers another pragmatic justification for calibrating ϕ.

[Figure 6 here]

One way of determining an appropriate value of ϕ is to model the time series ϕ_t as a constant with autoregressive residuals (the null hypothesis of a unit root in ϕ_t is strongly rejected); i.e $\phi_t = \phi + e_t$, with $e_t = \sum_{i=1}^{I} \rho_i e_{t-i} + u_t$. For

---

[16] The alternative of jointly estimating of all cross sections by the Newton-Raphson method while correcting for heteroscedasticity and non-normality, and with a varying number of instruments at each point in time, is practically infeasible.

$N = 3$, the point-estimate of ϕ is 1.08, with a standard deviation of 0.27.[17] This compares with the median value of ϕ_t of 0.91. The "middle ground" between these two estimates suggests that $\phi = 1$ is appropriate, and this hypothesis cannot be rejected by the t-statistic of $(1 - 1.08)/0.27 = 0.31$. The results for N ranging from 2 to 5 are similar.

Regarding the appropriate value for N, figure 7 illustrates summary statistics for the time series of the maximum absolute price residual from each individual cross section (associated with any instrument within that cross section), with N ranging from 1 to 5, and with $\phi = 1$ or ϕ_t for each cross section. It is evident that increasing N beyond 3 and and/or choosing $\phi = 1$ or ϕ_t makes little marginal contribution to the overall goodness of fit. For example, with $N = 3$ and $\phi = 1$, the full-sample maximum absolute price residual (i.e the maximum for the time series of the maximum absolute price residual from each individual cross section) is \$6,616. This is already well below 1 percent of the approximately \$1 million market-value for each of the instruments, and the marginal improvement through adding another mode is only \$929 (i.e the full-sample maximum absolute price residual with $N = 4$ and $\phi = 1$ is \$5,687), or \$1,009 by adding two modes. A similar conclusion is evident for the median of the time series of the maximum absolute price residual.

[Figure 7 here]

---

[17] After allowing for the effects of serial correlation on the estimated parameter variance. See Hamilton (1994) pages 610 to 612 for further details.

3.3 The time series of OLP$(3, 1)$ parameters

Having calibrated N and ϕ, the OLP$(3, 1)$ model of the yield curve is now used to "condense" each daily observation of market-quoted yields for the government-risk and bank-risk yield curves over the sample period into a convenient time series of four essential parameters.[18] These are illustrated in figures 8 and 9, and their intuitiveness is apparent. For example, the negative value of the Slope parameter around July 1998 indicates that the yield curve was sharply inverted at this point in time. The yield curve then steepened (the Slope parameter rose) to be strongly positive around July 1999, before flattening again (the Slope parameter fell) to be approximately flat by July 2000 (the Slope parameter was approximately zero). Similar generalisations may be made about the other parameters. In particular, the global widening of credit spreads during the 1998 Asian/Russian/Long Term Capital Management financial crisis is well captured by the increase in the fitted Swaps spread parameter at that time.

[Figure 8 here]

[Figure 9 here]

---

[18] There will also be a time series of price residuals associated with each instrument. The empirical results suggest that these are small enough to be ignored for the purposes of this article, but the issue is raised again in the conclusion.

4 The OLP$(3, 1)$ model under the expectations hypothesis of the yield curve

Having derived a time series of cross-sectional parameters that satisfactorily describe the yield curve at each point in time, the remainder of the article investigates the inter-temporal consistency of those parameters. In particular, given that the yield curve should naturally embody expectations of the evolution of interest rates, the expectations hypothesis of the yield curve (hereafter EH) is used as the basis for this investigation.

It is worth noting on the outset that the procedure that follows is not strictly dependent on the OLP$(3, 1)$ specification; this is just a convenient example. In general, defining N and L in the OLP(N, L) model sets the dimensions of the appropriate time-series model, which may then be derived analytically by following the approach in the text.[19]

4.1 The expectations hypothesis for the forward rate curve

The EH for the forward rate curve in continuous time may be written as:[20]

$$E_t \left[f \left(t + \tau, m \right) \right] = a \left(m \right) + f \left(t, m + \tau \right) \tag{12}$$

where:

---

[19]For example, the derived OLP$(4, 1)$ VAR specification is available from the author on request.

[20]The identity is derived by noting that under the EH, the forward rate is the expectation of the instantaneous interest rate $r(t)$ (for example, see Cochrane (2001) pages 352 to 355 for the discrete time analogue). Hence, $f \left(t, m + \tau \right) = E_t \left[r \left(t + m + \tau \right) \right] = E_t \left\{ E_{t+\tau} \left[r \left(t + m \right) \right] \right\} = E_t \left[f \left(t + \tau, m \right) \right]$.

- E_t is the expectation operator conditional upon information available at time t;

- τ is a positive increment of time, measured in years;

- $f(t+\tau, m)$ is the instantaneous rate m-years forward, measured at time $t+\tau$;

- $f(t, m+\tau)$ is the instantaneous rate $(m+\tau)$-years forward, measured at time t; and

- $a(m)$ allows for a term premium, which is a general function of maturity, but is strictly time-invariant in this analysis.

Therefore, under the EH, the initial shape of $f(t, m)$ at a point in time implies an expected evolution of $f(t, m)$ over time.

4.2 Application of the expectations hypothesis to the OLP$(3, 1)$ model

As at time t, the EH evolution of $f(t, m)$ as described by the OLP$(3, 1)$ model is calculated using equations 9 and 12, and an initial value of $\boldsymbol{\beta}(t)$ obtained from fitting the forward curve or the yield curve at time t. $E_t[f(t+\tau, m)]$ also has an OLP representation, i.e $E_t[f(t+\tau, m)] = E_t\left\{[\boldsymbol{\beta}(t+\tau)]'\mathbf{g}(\phi, m)\right\}$, and so the resulting equality is:

$$E_t\left\{[\boldsymbol{\beta}(t+\tau)]'\mathbf{g}(\phi, m)\right\} = \boldsymbol{\alpha}'\mathbf{g}(\phi, m) + [\boldsymbol{\beta}(t)]'\mathbf{g}(\phi, m+\tau) \qquad (13)$$

where $\boldsymbol{\alpha}'\mathbf{g}(\phi,m) = a(m)$, with both $\boldsymbol{\alpha}'$ and $\mathbf{g}(\phi,m)$ time-invariant. $\mathbf{g}(\phi,m+\tau)$ may be expressed precisely in terms of $\mathbf{g}(\phi,m),$[21] which means the EH evolution of the OLP$(3,1)$ model parameters can be written as:

$$E_t\left\{[\boldsymbol{\beta}(t+\tau)]'\right\}\mathbf{g}(\phi,m) = \boldsymbol{\alpha}'\mathbf{g}(\phi,m) + [\boldsymbol{\beta}(t)]'\left[\boldsymbol{\Phi}(\phi,\tau)\right]'\mathbf{g}(\phi,m)$$

where:

- $[\boldsymbol{\Phi}(\phi,\tau)]' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \exp(-\phi\tau) & 0 & 0 \\ 0 & -2\phi\tau\exp(-\phi\tau) & \exp(-\phi\tau) & 0 \\ 1-\exp(-\phi\tau) & 0 & 0 & \exp(-\phi\tau) \end{bmatrix}$.

Factoring out the common term $\mathbf{g}(\phi,m)$, and then taking the transpose gives the final result as:

$$E_t\left\{\boldsymbol{\beta}(t+\tau)\right\} = \boldsymbol{\alpha} + \boldsymbol{\Phi}(\phi,\tau)\boldsymbol{\beta}(t) \tag{14}$$

where $\boldsymbol{\beta}(t) = \left[\beta_1(t),\beta_2(t),\beta_3(t),\beta_{3+1}(t)\right]'$ is now a column vector of the OLP$(3,1)$ parameters. Hence, the EH evolution in continuous time of the forward rate or interest rate curve as described by the OLP$(3,1)$ model is conveniently summarised as a simple time-series process for the OLP$(3,1)$ model parameters. This process is denoted OLP$(3,1)/$EH hereafter.

---

[21] This is evident by re-expressing each $g_n(\phi,m+\tau)$ in terms of $g_n(\phi,m)$. For example, $g_2(\phi,m+\tau) = -\exp(-\phi[m+\tau]) = \exp(-\phi\tau)\cdot -\exp(-\phi m) = \exp(-\phi\tau)\cdot g_2(\phi,m)$. The other relevant results are: $g_1(\phi,m+\tau) = g_1(\phi,m)$; $g_3(\phi,m+\tau) = -2\phi\tau\exp(-\phi\tau)\cdot g_2(\phi,m) + \exp(-\phi\tau)\cdot g_3(\phi,m)$; and $g_{3+1}(\phi,m+\tau) = [1-\exp(-\phi\tau)]\cdot g_1(\phi,m) + \exp(-\phi\tau)\cdot g_{3+1}(\phi,m)$. Analogous results also follow for $N>3$.

4.3 A time-series model for the OLP$(3, 1)$ parameters

The continuous time result provides the basis for an iterated discrete time version of OLP$(3, 1)$/EH; i.e the evolution of $\boldsymbol{\beta}(t)$ in continuous time may be modeled as a finite difference process for "units of time" τ. Defining steps of τ with an integer index, and allowing for model residuals gives:

$$\boldsymbol{\beta}_t = \boldsymbol{\alpha} + \boldsymbol{\Phi}(\boldsymbol{\phi}, \boldsymbol{\tau})\, \boldsymbol{\beta}_{t-1} + \boldsymbol{\varepsilon}_t(\tau) \tag{15}$$

where:

- t as a subscript is now an integer index representing individual time-steps of τ; and

- $\boldsymbol{\varepsilon}_t(\tau)$ is an iid normal 4-vector.

In a forecasting sense, $\boldsymbol{\varepsilon}_t(\tau)$ represents the "expectational error" of $\boldsymbol{\beta}_t$ relative to $\boldsymbol{\beta}_{t-1}$. This is attributable to new information arriving continuously between time $t - \tau$ and t; that new information simultaneously causes the expectations formed at time $t-1$ to be incorrect, and changes future expectations which are now incorporated in the updated $\boldsymbol{\beta}(t)$. Writing the expectational error as a function of τ recognises that $\boldsymbol{\varepsilon}_t(\tau)$ should increase in magnitude for an increasing horizon τ, although a functional specification is not actually required for the empirical estimation.

In principle then, a time series of OLP$(3, 1)$ parameters might be expected to conform to a vector auto-regression (VAR) in levels, which could be tested empirically. In practice, however, the coefficient of 1 in the top-left entry of

$\boldsymbol{\Phi}(\phi,\tau)$ (associated with $\beta_{1,t}$) suggests that the VAR in levels is expected to contain a single unit root,[22] and hence empirical parameter estimates and tests of restrictions would have non-standard distributions.

Since there is no dependence of any parameter on the level of the Level parameter, an OLP$(3,1)$/EH VAR representation that conforms with both empirical and theoretical considerations is:

$$\boldsymbol{\lambda}_t = \boldsymbol{\gamma} + \boldsymbol{\Lambda}(\boldsymbol{\phi},\boldsymbol{\tau})\,\boldsymbol{\lambda}_{t-1} + \boldsymbol{\varepsilon}_t(\tau) \tag{16}$$

where:

- $\boldsymbol{\gamma} = [\delta, \alpha_2, \alpha_3, \alpha_{3+1}]'$;

- $\boldsymbol{\lambda}_t = \left[\Delta\beta_{1,t}, \beta_{2,t}, \beta_{3,t}, \beta_{3+1,t}\right]'$; and

- $\boldsymbol{\Lambda}(\phi,\tau) = \begin{bmatrix} 0 & 0 & 0 & 1-\exp(-\phi\tau) \\ 0 & \exp(-\phi\tau) & -2\phi\tau\exp(-\phi\tau) & 0 \\ 0 & 0 & \exp(-\phi\tau) & 0 \\ 0 & 0 & 0 & \exp(-\phi\tau) \end{bmatrix}$

An equivalent option is to calculate and estimate a complete model in first differences on the left-hand side (i.e using $\Delta\boldsymbol{\beta}_t = \boldsymbol{\Phi}(\boldsymbol{\phi},\boldsymbol{\tau})\boldsymbol{\beta}_{t-1} - \boldsymbol{\beta}_{t-1} = [\boldsymbol{\Phi}(\boldsymbol{\phi},\boldsymbol{\tau}) - \mathbf{I}]\boldsymbol{\beta}_{t-1}$), although this no longer has the convenient VAR form of equation 16. However, the four individual equations from the first difference specification do provide convenient single-equation tests of each OLP$(3,1)$ component under the EH, i.e:

---

[22] The eigenvalues of $\boldsymbol{\Phi}(\phi,\tau)$ are $\{1, \exp(-\phi\tau), \exp(-\phi\tau), \exp(-\phi\tau)\}$, and $|\exp(-\phi\tau)|$ is strictly less than 1.

$$\Delta\beta_{1,t} = \eta_1 + \theta_1 \cdot \left\{\beta_{3+1,t-1} \cdot [1 - \exp\left(-\phi\tau\right)]\right\} + \varepsilon_{1,t} \qquad (17\text{a})$$

$$\Delta\beta_{2,t} = \eta_2 + \theta_2 \cdot \left\{ \begin{array}{l} \beta_{2,t-1} \cdot [\exp\left(-\phi\tau\right) - 1] \\[2ex] -\beta_{3,t-1} \cdot 2\phi\tau\exp\left(-\phi\tau\right) \end{array} \right\} + \varepsilon_{2,t} \qquad (17\text{b})$$

$$\Delta\beta_{3,t} = \eta_3 + \theta_3 \cdot \left\{\beta_{3,t-1} \cdot [\exp\left(-\phi\tau\right) - 1]\right\} + \varepsilon_{3,t} \qquad (17\text{c})$$

$$\Delta\beta_{3+1,t} = \eta_{3+1} + \theta_{3+1} \cdot \left\{\beta_{3+1,t-1} \cdot [\exp\left(-\phi\tau\right) - 1]\right\} + \varepsilon_{3+1,t} \qquad (17\text{d})$$

If the time series of each component conforms to OLP$(3,1)$/EH, then equations 17a to 17d should result in estimates of θ_1 to θ_{3+1} that are equal to 1. These individual equations also provide some quantitative intuition to the stylised facts of mean reversion in the shape of the yield curve; for example, if the yield curve is steeply sloped (i.e $\beta_{2,t-1} \gg 0$), then equation 17b predicts that the yield curve is more likely to flatten in the future (i.e $\Delta\beta_{2,t} < 0$, since $[\exp\left(-\phi\tau\right) - 1] < 0$). This tendency to flatten will also be influenced by the extent that the yield curve is "upwardly bowed" (as measured by $\beta_{3,t} > 0$). Similar generalisations may be made for the expected changes to the other parameters.

5 Empirical tests of New Zealand OLP$(3,1)$ data against the EH predictions

This section tests the New Zealand OLP$(3,1)$ time-series data derived in section 3 against the OLP$(3,1)$/EH predictions derived in section 4. The results

presented are only for the weekly horizon (i.e $\tau = 7/365$ years); the daily and monthly horizon results were similar in all respects. To generate the weekly horizon data set, the OLP(3, 1) parameters are sampled each Wednesday, or the prior or closest trading day in the event of a market holiday. This selection avoids the unnecessary complication of having to model the moving-average process that would be introduced by using data with overlapping horizons, and the span of the data still allows for 263 weekly observations.

5.1 Unit root and cointegration results

Figure 10 shows the results of Augmented Dickey-Fuller (ADF) tests on the first differences and the level of each time series (note that in this and each of the following figures *, **, and *** respectively represent rejection of the null hypothesis at the 10, 5, and 1 percent level of significance). The null hypothesis of a unit root in the first differences is rejected in each case. The results for the levels are generally in agreement with the OLP(3, 1)/EH predictions, in that a unit root cannot be rejected for the Level series, and can be rejected for the Slope and Bow series.

The exception is the Swaps spread series, where a unit root cannot be rejected at traditional levels of significance. This result is probably sample specific, given that the data period includes the sharp moves resulting from the 1998 global financial crisis. Further, an Engle-Granger test for cointegration between the Level and Swaps spread series strongly rejects the hypothesis of cointegration (regardless of ordering), which is consistent with the VAR specification in equation 16. Hence, the estimation and hypothesis testing of the

26

VAR in equation 16 proceeds on the assumption that the Swaps spread series does not truly contain a unit root.

[Figure 10 here]

5.2   Vector auto-regression results

Figure 11 compares the coefficient matrix obtained from estimating the VAR specified in equation 16 with the coefficient matrix predicted by the $\text{OLP}(3, 1)/\text{EH}$ (i.e using $\tau = 7/365$ years in equation 16), and assuming a constant vector of zero (i.e the pure EH). Treating this as an unrestricted and fully-restricted estimation, the likelihood ratio statistic has a probability value of 6.7 percent, which suggests that the data do not reject the $\text{OLP}(3, 1)/\text{EH}$ model at the 5 percent level of significance.

[Figure 11 here]

However, there are obviously many alternative forms that the estimated and/or fully-restricted VARs could take, and so figure 12 shows the results of a systematic series of combinations. Firstly, the data do not reject a first-order VAR against a second-order VAR, suggesting that the estimated first-order VAR captures the dynamics of the data adequately. Secondly, allowing for a positive term premium in the Swaps spread parameter, i.e using $\alpha_{3+1} = 1.6$ basis points (Case 1), or $\delta = -1.6$ basis points and $\alpha_{3+1} = 1.6$ basis points (Case 2), notably improves the fit of the $\text{OLP}(3, 1)/\text{EH}$ model to the data.[23] This suggests that

---

[23]The value of 1.6 basis points is determined using the average of 85 basis points for the Swaps spread series over the data period, and noting that, according to the $\text{OLP}(3, 1)/\text{EH}$, the Swaps spread should follow an AR(1) process with a coefficient of $\rho = \exp(-7/365)$ for the weekly horizon. Equating 85 basis points to the expected long-run value for the AR1 process, $\alpha_{3+1}/(1 - \rho)$, gives $\alpha_{3+1} = 85 \times (1 - \rho) = 1.6$. Using -1.6 basis points as the constant associated with the Level first difference means that a Swaps spread at its average level of 85

a model with a positive term premium in the bank-risk yield curve is "more realistic" than assuming the pure EH (which is not surprising given that Swaps spreads tend to remain exclusively positive in the yield curves of almost all countries, rather than reverting to zero). Note that a term premium in the base yield curve could be allowed for using $\alpha_2 > 0$, but the empirical results suggests this is not needed at this stage.

[Figure 12 here]

5.3 Single equation model results

Figure 13 shows the results from individually estimating the equations 17a to 17d. The initial results from the OLS estimations generally showed evidence of non-normality and heteroscedasticity in the residuals, as summarised in figure 13. Given the potential ambiguity in the source and structure of heteroscedasticity (the evidence is mixed between autoregressive conditional heteroscedasticity and/or a structural change in volatility[24]), the standard errors of the OLS estimates are therefore adjusted using White's heteroscedasticity-consistent estimator (or the Newey-West method for Swaps spread equation, given the evidence of serial correlation in the residuals). These results are shown in the bottom three rows of figure 13.

Apart from the Swaps spread equation, the results do not reject the OLP$(3, 1)/$ pure EH predictions (i.e η_1 to $\eta_{3+1} = 0$), at the standard levels of significance.

---

basis points will have no tendency to increase or decrease the Level parameter over the next period.

[24]Prior to the Official Cash Rate regime, the Reserve Bank used a Monetary Conditions Index (or MCI - a simple combination of short-term interest rates and the exchange rate) to indicate the desired stance of monetary policy. The adherence to an MCI sometimes led to sharp changes in short-term interest rates due to changes in the exchange rate.

However, the results for both the Level and Swaps spread equations are again notably improved by setting $\eta_1 = -1.6$ basis points and $\eta_{3+1} = 1.6$ basis points, as discussed in the previous section.

[Figure 13 here]

6   Conclusion and areas for further work

The empirical application of the OLP model to New Zealand data indicates that daily government-risk and bank-risk yield curve data is adequately represented using three parameters for the base yield curve, and one additional spread parameter, i.e by the $OLP(3,1)$ model. The predictions of the pure expectations hypothesis applied to the $OLP(3,1)$ model are marginally accepted by New Zealand data, although further results suggest that allowing for a term premium in the bank-risk yield curve provides a more acceptable representation.

An obvious extension of the empirical work in this article is to apply the OLP model to the yield curves of other countries. Apart from providing similar information to that presented here for New Zealand, this may also provide a basis for modelling the transmission of yield curve effects between countries.

In general, the dual cross-sectional and inter-temporal consistency of the $OLP(3,1)$ model makes it applicable to a variety of yield curve related topics. For example, one application is to provide a gauge of interest rate expectations held by the market at any point in time, and to measure changes in those expectations over time. A related application is to predict the evolution of the yield curve, which may be useful for market trading. Another aspect that may

be applicable to market trading is the systematic analysis of the price residuals that arise as a by-product of each cross-sectional fitting of the yield curve.

Regarding the OLP model itself, it would be desirable to provide more robust theoretical foundations for the proposed modal formulation, and also to investigate the flexibility of the model to alternative forms of expectations hypothesis, such as risk-neutral relationships.

All of the extensions noted above are currently being investigated by the author.

References

Bank for International Settlements 1999, *Zero-coupon yield curves: technical documentation*, Bank for International Settlements, Switzerland.

Barrett, W, T Gosnell and A Heuson 1995, 'Yield curve shifts and the selection of immunization strategies', *Journal of Portfolio Management* **Fall**, 53–64.

Bliss, R 1997, 'Testing term structure estimation methods', *Advances in Futures and Options Research* **9**, 197–231.

Cochrane, J 2001, *Asset Pricing*, Princeton University Press.

Courant, R and D Hilbert 1953, *Methods of Mathematical Physics, First English Edition*, Interscience Publishers Inc., New York.

Dahlquist, M and L Svensson 1994, 'Estimation of the term structure of interest rates with simple and complex functional forms: Nelson and Siegel vs. Longstaff and Schwartz', *IIES Seminar Paper* **565**.

Diebold, F and C Li 2002, 'Forecasting the term structure of government bond yields', *NBER Working Paper* .

Echols, M and J Elliot 1976, 'Parsimonious modelling of yield curves', *Journal of Financial and Quantitative Analysis* **October**, 87–114.

Fergusson, R and S Raymar 1998, 'A comparative analysis of several popular term structure estimation models', *Journal of Fixed Income* **March**, 17–32.

Hamilton, J 1994, *Time Series Analysis*, Princeton University Press.

Helwege, J and C Turner 1999, 'The slope of the credit yield curve for speculative-grade issuers', *Journal of Finance* **54**, 1869–1884.

Hördahl, P 2000, 'Estimating the implied distribution of the future short term interest rate using the Longstaff-Schwartz model', *European Central Bank Working Paper Series* **16**.

Houweling, P, J Hoek and F Kleibergen 2001, 'The joint estimation of term structures and credit spreads', *Journal of Empirical Finance* **8**, 297–323.

Hunt, B 1995, 'Fitting parsimonious yield curve models to Australian coupon bond data', *University of Technology Sydney, School of Finance and Economics Working Paper No. 51* .

Jarrow, R, D Lando and S Turnbull 1997, 'A Markov model for the term structure of credit risk spreads', *Review of Financial Studies* **10(2)**, 481–523.

Kacala, V 1993, 'Zero coupon curves part 3: Fair bond yields', *Buttle Wilson* .

Lin, B 2002, 'Fitting term structure of interest rates using B-splines: the case of Taiwanese Government bonds', *Applied Financial Economics* **12**, 57–75.

Mansi, S and J Phillips 2001, 'Modeling the term structure from the on-the-run Treasury yield curve', *The Journal of Financial Research* **24(4)**, 545–564.

Nelson, C and A Siegel 1987, 'Parsimonious modelling of yield curves', *Journal of Business* **October**, 473–489.

Pham, T 1998, 'Estimation of the term structure of interest rates: an international perspective', *Journal of Multinational Financial Management* **8**, 265–283.

Rainville, E and P Bedient 1981, *Elementary Differential Equations - 6th ed.*, Collier Macmillan Publishers.

Schich, S 1997, 'Estimating the German term structure', *Deutsche Bundesbank, Economic Research Group, Discussion Paper* **4 (October)**.

Seppala, J and P Viertio 1996, 'The term structure of interest rates: estimation and interpretation', *Bank of Finland Discussion Papers* **19/96**.

Söderlind, P and L Svensson 1997, 'New techniques to extract market expectations from financial instruments', *Journal of Monetary Economics* **40**, 383–429.

Svensson, L 1994, 'Estimating and interpreting forward interest rates: Sweden 1992-4', *Centre for Economic Policy Research* **Discussion paper 1051**.
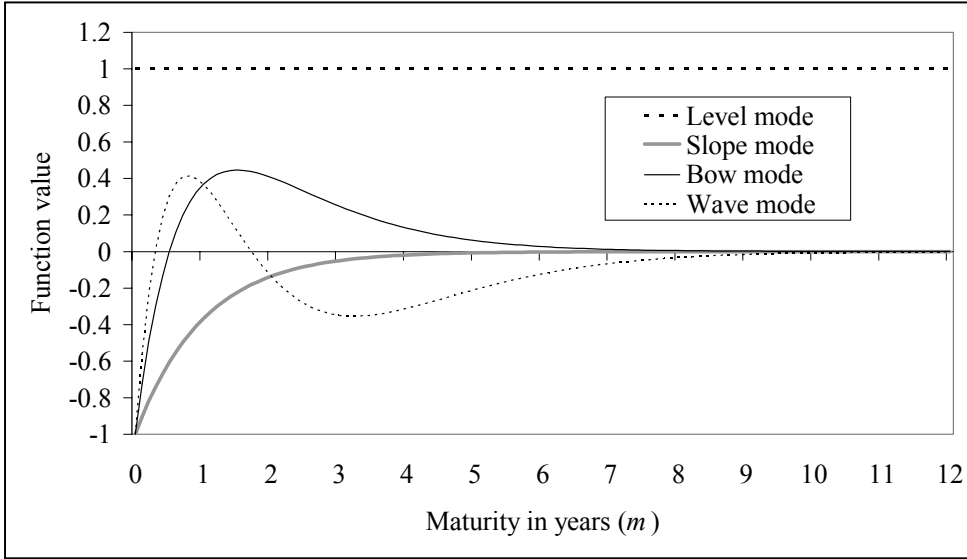
Figure 1: The first four forward rate modes of the OLP model. Level mode is g_1, Slope mode is g_2, Bow mode is g_3, and Wave mode is g_4, all with $\phi = 1$.
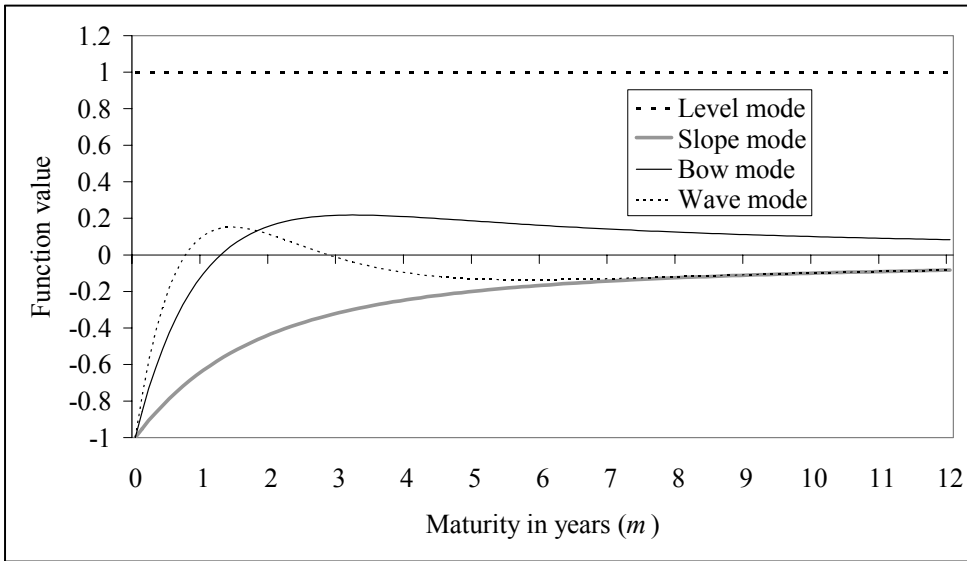

Figure 2: The first four interest rate modes of the OLP model. Level mode is s_1, Slope mode is s_2, Bow mode is s_3, and Wave mode is s_4, all with $\phi = 1$.
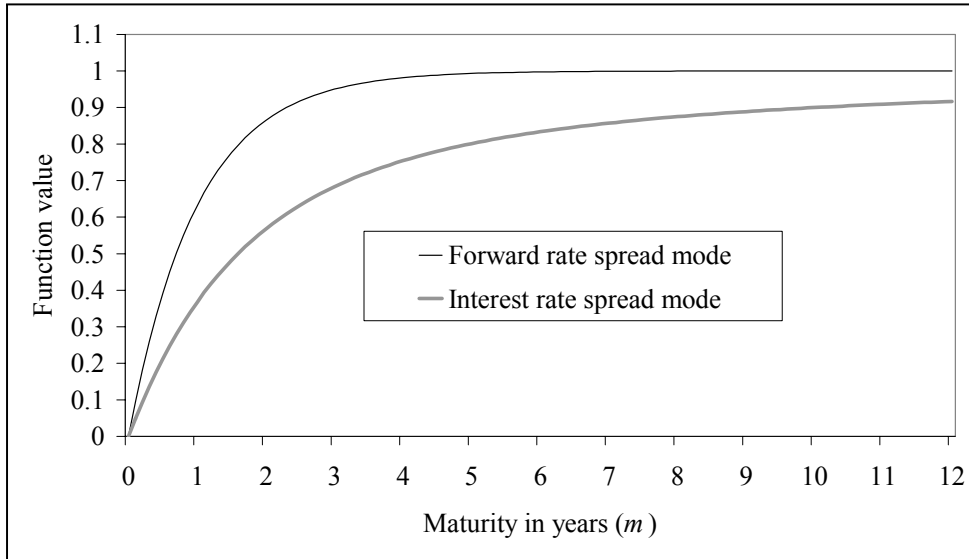
Figure 3: The forward rate spread mode and the interest rate spread mode for the OLP model, both with $\phi = 1$.
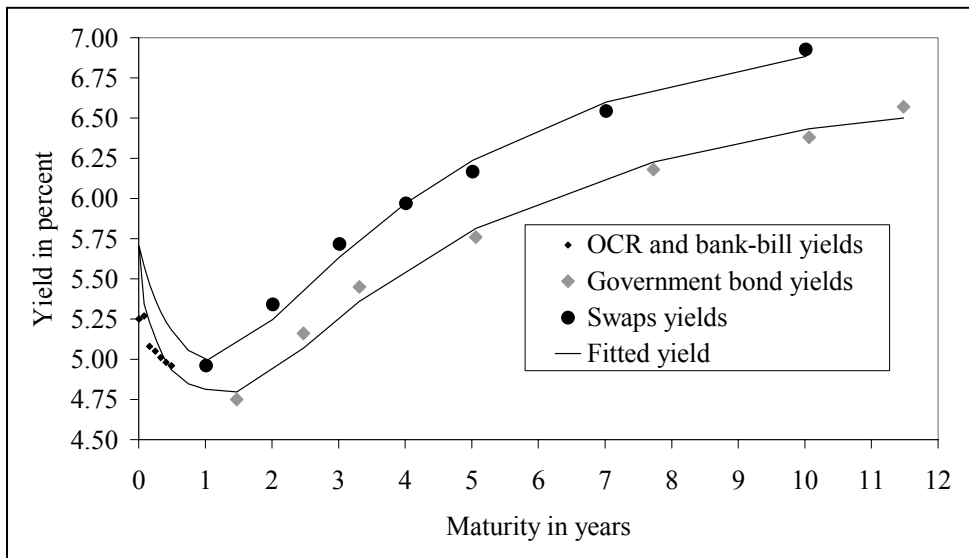

Figure 4: Market yields for 25 October 2001, and the estimated OLP$(3, 1)$ yield curve with $\phi = 1$. The estimated parameter values (in percent) are Level = 7.17 (β_1), Slope = +4.40 (β_2), Bow = -2.65 (β_3), and Swaps spread = 0.52 (β_{3+1}).
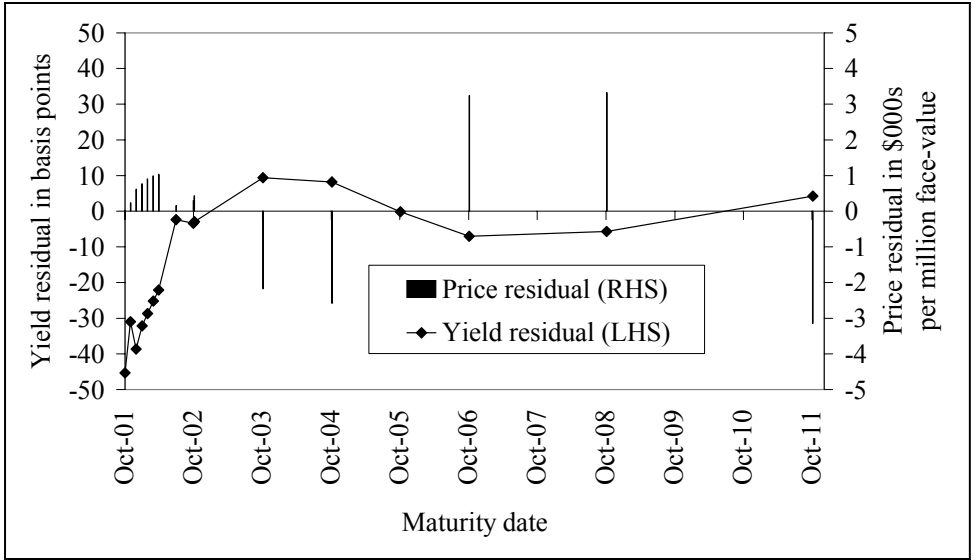
Figure 5: Yield and price residuals associated with the 25 October 2001 estimated OLP(3, 1) yield curve with $\phi = 1$, as in figure 4. Government bond residuals have been omitted for clarity.
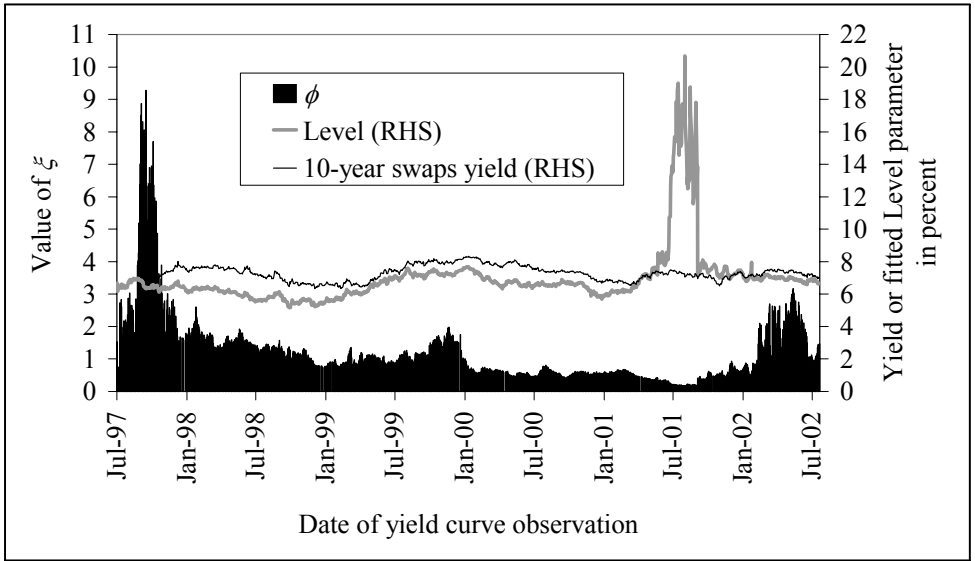

Figure 6: Time series of ϕ_t (i.e ϕ variable for each cross section of the yield curve) and Level parameter estimates for the OLP(4, 1) model.
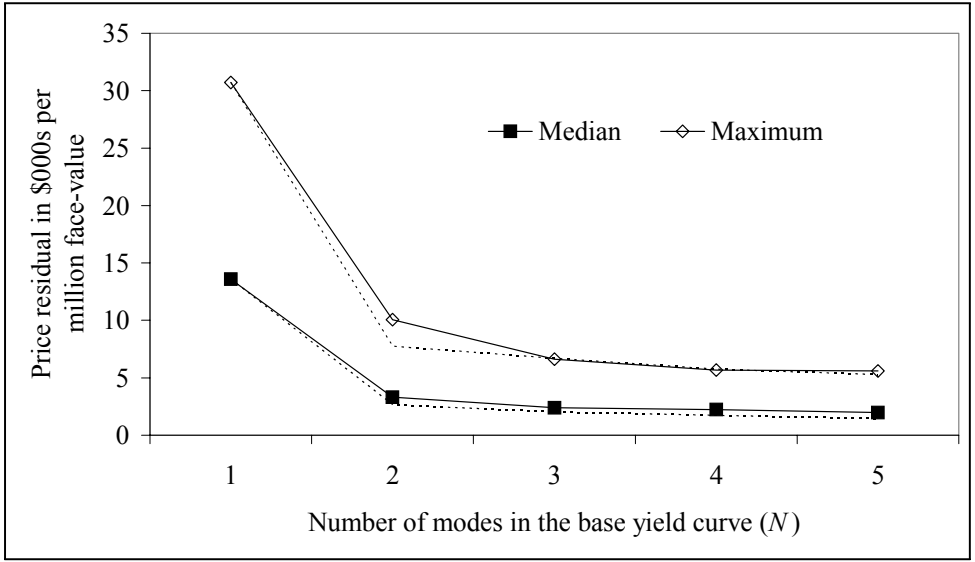
Figure 7: Summary statistics for the time series of the maximum absolute price residual from each individual cross section using the OLP(N, 1) model with $\phi = 1$. N ranges from 1 to 5, and the associated dotted lines use ϕ_t instead of $\phi = 1$.

Figure 8: Time series of the estimated OLP(3, 1) Level (β_1) and Swaps spread (β_{3+1}) parameters. The scale range for Level is identical to that for Slope and Bow in figure 11, but the Swaps spread scale range is only 2 percentage points.

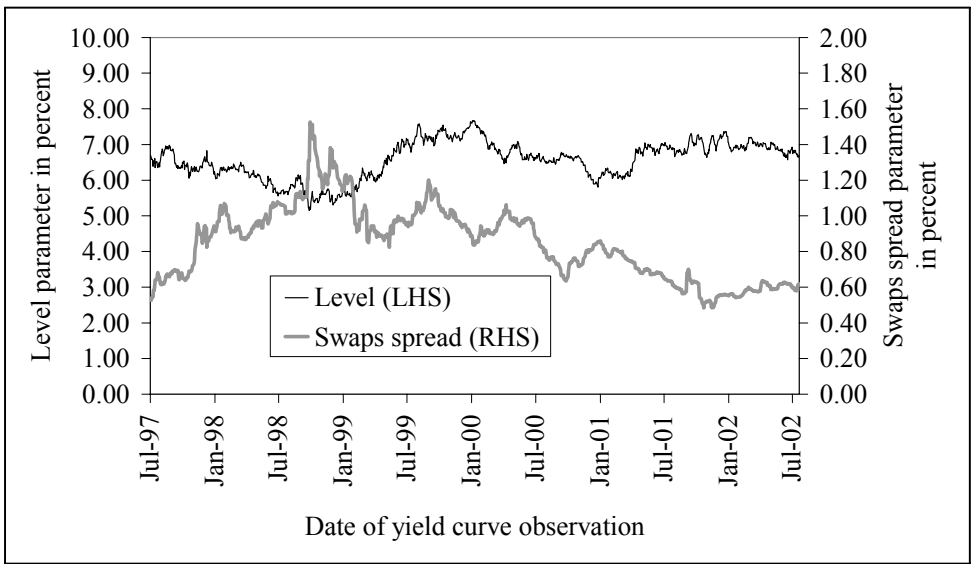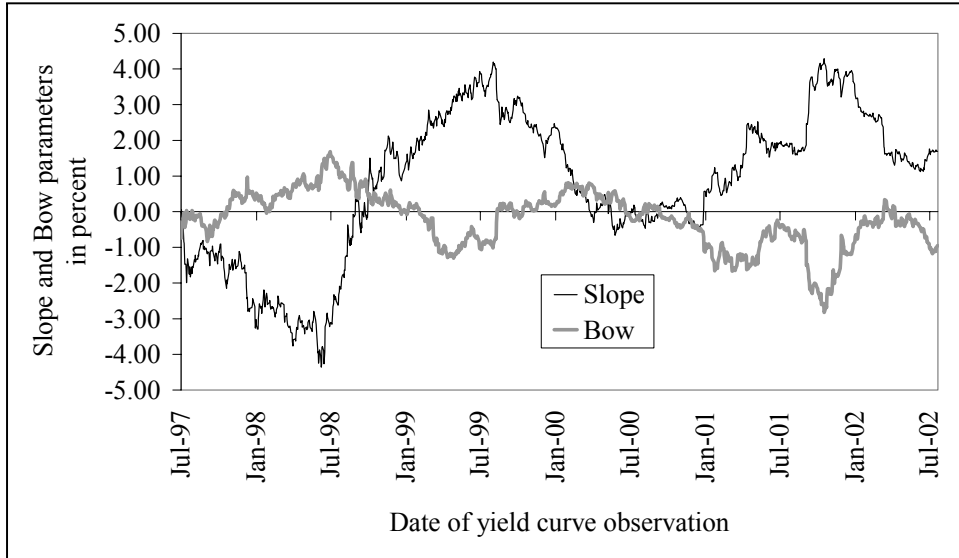Figure 9: Time-series of the estimated OLP$(3,1)$ Slope (β_2) and Bow (β_3) parameters.

ADF unit root tests	Number of lags	ADF statistic
Level, no constant	15	0.06
Slope, no constant	15	-2.01**
Bow, no constant	15	-2.07**
Swaps spread, no constant	15	-0.48
Level, with constant	0	-2.20
Swaps spread, with constant	1	-1.86
ΔLevel, no constant	0	-16.52***
ΔSlope, no constant	0	-16.45***
ΔBow, no constant	1	-15.66***
ΔSwaps spread, no constant	0	-13.92***

Figure 10: ADF tests for unit roots on each time series of the OLP$(3,1)$ model parameters. Sample size is 263 for levels, and 262 for first differences. Critical values used to test significance were sourced from Hamilton (1994) page 763.

Unrestricted	Constant	ΔLevel	Slope	Bow	Swaps
ΔLevel	-7.7	-0.048	0.001	-0.026	0.085
Slope	-16.9	-0.286	0.974	-0.074	0.223
Bow	-1.7	0.067	-0.003	0.963	0.009
Swaps spread	3.8	-0.020	0.000	0.011	0.957
Twice log likelihood 5479.55, 241 degrees of freedom.					
Fully restricted	Constant	ΔLevel	Slope	Bow	Swaps
ΔLevel	0	0	0	0	0.019
Slope	0	0	0.981	-0.038	0
Bow	0	0	0	0.981	0
Swaps spread	0	0	0	0	0.981
Twice log likelihood 5449.38, 20 restrictions.					

Figure 11: First-order VAR estimate against the pure EH. The top entries represent the estimated constant vector (expressed in basis points) and 4×4 coefficient matrix. The bottom table entries represent $\boldsymbol{\gamma} = \mathbf{0}$ and the OLP$(3, 1)$/EH 4×4 coefficient matrix $\boldsymbol{\Lambda}$. The likelihood ratio statistic is 30.17, and $\chi^2 (30.17)$ with 20 degrees of freedom has a probability value of 6.7 percent.

Estimated VAR	Restricted VAR (see notes)	Likelihood ratio statistic	Number of restrictions	Probability value (percent)
Second-order	First-order	12.67	16	69.7
First-order	Pure EH	30.17	20	6.7*
First-order	EH, Case 1	16.23	20	70.3
First-order	EH, Case 2	12.59	20	89.4

Figure 12: First- and second-order VAR estimates against various alternatives. The pure EH uses $\boldsymbol{\gamma} = \mathbf{0}$, Case 1 uses $\boldsymbol{\gamma} = (0, 0, 0, 1.6)'$ basis points, and Case 2 uses $\boldsymbol{\gamma} = (-1.6, 0, 0, 1.6)'$ basis points (as discussed in the text).

	θ_1	θ_2	θ_3	θ_4	θ_1	θ_4
OLS coefficient	0.12	0.84	1.51	0.06	1.73	1.17
Standard deviation	0.54	0.66	0.82	0.16	2.28	0.67
Jarque-Bera test	0.14	78.42	54.45	2004.49	0.10	2415.05
p value	93.4	0.0***	0.0***	0.0***	95.2	0.0***
Ljung-Box test (3 lags)	1.75	1.42	1.27	13.13	1.62	12.67
p value	62.5	70.2	73.5	0.4***	65.6	0.5***
ARCH test (3 lags)	11.29	15.01	7.16	4.15	12.73	3.12
p value	1.0**	0.2***	6.7*	24.6	0.5***	37.3
Goldfeld-Quandt test	1.20	2.73	0.91	5.25	1.20	5.11
p value (two sided)	15.2	0.0***	68.3	0.0***	15.2	0.0
Adjusted standard deviation	0.54	0.62	0.99	0.19	2.43	0.91
H0: α-1=0 t-statistic	-1.63	-0.26	0.51	-4.92	0.30	0.19
p value (two sided)	10.5	79.8	60.8	0.0***	76.4	84.9

Figure 13: Single equation estimations, and tests against the relevant OLP$(3,1)$/pure EH predictions for θ. The right-hand columns include $\eta_1 = -1.6$ basis points and $\eta_{3+1} = 1.6$ basis points. All test significances are reported as probability values in percentages, allowing for the appropriate degrees of freedom.