Working Paper 99-18 / Document de travail 99-18

Estimating One-Factor Models of Short-Term Interest Rates

by

Des Mc Manus and David Watt

Bank of Canada Banque du Canada

ISSN 1192-5434
ISBN 0-662-28308-2

Printed in Canada on recycled paper

Bank of Canada Working Paper 99-18

November 1999

Estimating One-Factor Models of Short-Term Interest Rates

by

Des Mc Manus and David Watt*

Financial Markets Department
Bank of Canada
Ottawa, Canada K1A 0G9
E-mail: dmcmanus@bank-banque-canada.ca
dwatt@cmhc-schl.gc.ca

*Current address: Canada Mortgage and Housing Corporation, Ottawa, Canada.

Contents

Acknowledgements

The authors would like to thank Ben Fung, Paul Gilbert, Peter Thurlow, Mark Zelmer, and participants at the 1999 Northern Finance Association conference for their helpful comments. Thanks also to Ronald Gallant and George Tauchen for a base set of computer programs.

Abstract

There currently exists in the literature several continuous-time one-factor models for short-term interest rates. This paper considers a wide range of these models that are nested into one general model. These models are approximated using both a discrete-time model and a model that accounts for aggregation effects over time, and are estimated by both the method of maximum likelihood and the general method of moments, for both Canadian and U.S. data. The estimation results are found to be independent of the approximation model used. However, the results are dependent on the estimation technique, more so for Canada than the United States. As an alternative check, the efficient method of moments is also employed. Hypothesis testing strongly suggests these one-factor models do not provide a good description of the evolution of Canadian short-term interest rates. In contrast, these models perform better for short-term U.S. interest rates.

JEL classification: C52, G10

Bank of Canada classification: Financial markets; Interest rates

Résumé

On relève plusieurs modèles à un facteur formulés en temps continu dans les ouvrages économiques pour décrire le comportement des taux d'intérêt à court terme. Les auteurs de l'étude examinent une large gamme de ces modèles constituant des cas particuliers d'un modèle plus général. Ils les représentent de façon approchée au moyen d'un premier modèle en temps discret et d'un second qui tient compte des effets d'agrégation au fil du temps, qu'ils estiment ensuite par la méthode du maximum de vraisemblance et la méthode des moments généralisés à l'aide de données canadiennes et américaines. Les résultats de l'estimation ne varient pas selon l'approximation utilisée; ils varient toutefois en fonction de la méthode d'estimation, et ce davantage dans le cas du Canada que pour les États-Unis. Les auteurs emploient aussi la méthode des moments efficaces pour contre-vérifier leurs résultats. Les tests d'hypothèse donnent fortement à penser que les modèles à un facteur parviennent mal à décrire l'évolution des taux à court terme au Canada. Ces modèles arrivent mieux, en revanche, à expliquer celle des taux comparables aux États-Unis.

Classification JEL: C52, G10

Classification de la Banque du Canada: Marchés financiers; Taux d'intérêt

1. Introduction and overview

The short-term interest rate is important in many financial economics models, such as models of the term structure of interest rates, bond pricing models, and derivative security pricing models. Short-term interest rates are also important in the development of tools for effective risk management and in many empirical studies analyzing term premiums and yield curves, where risk-free short-term interest rates are taken as reference rates for other interest rates.

Short-term interest rates are also a crucial feature of the monetary transmission mechanism. Duguay (1994) describes the monetary transmission mechanism as starting with a monetary authority's actions influencing short-term interest rates and the exchange rate, which then go on to ultimately affect aggregate demand and inflation. Thus, to fully characterize the monetary transmission mechanism, it is imperative to have a good model of the behaviour of short-term interest rates.

As a first step in modelling short-term interest rates, one-factor models of the term structure of interest rates will be discussed and their applicability to Canada analyzed. These models are the basic building blocks for more complicated models of the term structure where the short-term interest rate represents the single factor. Thus, finding an adequate characterization of the short-term interest rate will help determine if one-factor models of the term structure may be gainfully applied to Canadian interest rates. If these models prove insufficient, the empirical analysis may indicate alternate paths of investigation for other factors that could characterize the term structure of interest rates in Canada.

Several models have been proposed for short-term interest rates, but until relatively recently, they had not been formally compared. Chan, Karolyi, Longstaff, and Sanders (1992) (hereafter CKLS) estimated and compared several models of short-term interest rates to explain U.S. 1-month Treasury bill yields. The results indicated that models that allowed the variability of interest rates to depend upon the level of interest rates captured the dynamic behaviour of short-term interest rates more successfully. The level effect was such that interest rate volatility was positively correlated with the level of interest rates.

Tse (1995) and Dahlquist (1996) extended the analysis of CKLS to international short-term interest rates.Their results indicated that in many countries the impact of the level of rates upon the volatility of interest rates was also positive, though lower than in the United States. Tse (1995) found that the impact of the level of interest rates on volatility was negative for Canada, the only country where a negative impact was discovered. However, since the impact parameter was not statistically significant, Tse did not discuss the result in detail.

The main goal of this paper is to determine if Canadian short-term interest rates can be adequately modelled using a one-factor model. For comparative purposes, the appropriateness of one-factor models for the U.S. short-term interest rate is also investigated. Attention is focused on the class of one-factor models proposed by CKLS that includes a wide range of notable one-factor models, though the class does not encompass all possible one-factor models.

The analysis of the CKLS models consists of choosing an analytic expression for the evolution of the short-term interest rate and an estimation technique. Two alternative, but related, analytic expressions are considered. The first is a discrete-time approximation to the continuous-time model. The second is an alternative discrete-time model that is formulated to reduce potential temporal aggregation bias from the discretization of the continuous-time process. There are also several techniques for estimating one-factor models. The above observations prompt several key questions—which are addressed in this paper. Do the estimation results depend on the choice of analytic expression for the one-factor models? Are the results sensitive to the estimation techniques employed? And do the results depend on the data set analyzed? Are the results for Canada and the United States different?

The analysis indicates that the estimation results tend to be independent of the analytic approximation used to characterize the one-factor model. The discrete-time approach yields estimates that are almost identical to the more exact time-aggregation approach, owing to the relatively minor degree of mean reversion in short-term interest rates.

The estimation results are found to depend on both the country under consideration and the estimation technique employed.[1] The results for Canadian short-term interest rates are unstable and differ quite considerably depending on the estimation technique employed. Evidence suggests that the evolution of Canadian short-term interest rates cannot be adequately described by the CKLS class of one-factor models. On the other hand, the results for U.S. short-term interest rates are fairly stable. One-factor models do a much better job of describing U.S. short-term interest rate data than similar Canadian data—although weak evidence against one-factor models for U.S. short-term interest rate data appears to exist.

1. The question of whether the estimation results are independent of the time frame of the data is not investigated in this paper.

The paper is organized as follows: Section 2 presents the CKLS one-factor models for short-term interest rates, describing both the discrete-time and aggregation models. Section 3 reviews standard estimation techniques, namely the maximum likelihood method and the general method of moments. Section 4 discusses the data and the initial empirical results obtained by using both the maximum likelihood method and the general method of moments estimation techniques. Sections 5 and 6 review the efficient method of moments, an alternative estimation technique. Section 7 presents the results of the efficient method of moments estimation. Section 8 concludes the paper and discusses possible further work.

2. The model

CLKS proposed the following general model for short-term interest rates:

$$dr(t) = [\alpha + \beta r(t)] dt + \sigma r^{\gamma} dz \quad , \tag{1}$$

where r is the short-term interest rate and z is a geometric Brownian motion process. Thus, both the drift, $\alpha + \beta r(t)$, and the conditional variance of the interest rate process, $\sigma^2 r^{2\gamma} dt$, depend upon the level of the interest rate. Several well-known one-factor models can be derived from the above model through parametric restrictions. They are presented in the following table:

Model	**Specification**	**Restrictions**
Merton (1973)	$dr = \alpha dt + \sigma dz$	$\beta = 0, \gamma = 0$
Vasicek (1977)	$dr = (\alpha + \beta r) dt + \sigma dz$	$\gamma = 0$
Cox, Ingersoll, and Ross (1985)	$dr = (\alpha + \beta r) dt + \sigma r^{1/2} dz$	$\gamma = 1/2$
Dothan (1978)	$dr = \sigma r dz$	$\alpha = \beta = 0, \gamma = 0$
Geometric Brownian Motion	$dr = \beta r dt + \sigma r dz$	$\alpha = 0 \;,\; \gamma = 1$
Brennan and Schwartz (1980)	$dr = (\alpha + \beta r) dt + \sigma r dz$	$\gamma = 1$
Cox, Ingersoll, and Ross (1980)	$dr = \sigma r^{3/2} dz$	$\alpha = \beta = 0, \gamma = 3/2$
Cox (1975)	$dr = \beta r dt + \sigma r^{\gamma} dz$	$\alpha = 0$

The Merton (1973) model is a simple Brownian motion for short-term interest rates. The model of Vasicek (1977) is an Ornstein-Uhlenbeck process. The model of Cox, Ingersoll, and Ross (1985) is frequently referred to as the square-root process (CIR-SR). The Geometric Brownian Motion (GBM) was used by Black and Scholes (1973) to derive the prices of options, where σ was referred to as the implied volatility of the option. The model of Cox (1975) is often referred to as the Constant Elasticity of Variance (CEV) model. From the parametric restrictions, it is obvious that the models cannot generally be written as special cases of one another. That is, although each of the models is nested within (1), they are typically non-nested with respect to each other.

Typically the continuous-time model, (1), is discretized as follows:

$$r(t+1) - r(t) = \alpha + \beta r(t) + \varepsilon(t+1) \tag{2}$$

$$\text{where } E_t[\varepsilon(t+1)] = 0 \text{ and } E_t[\varepsilon^2(t+1)] = \sigma^2 r(t)^{2\gamma}. \tag{3}$$

The parameters of the model are then estimated using either maximum likelihood methods, for example Nowman (1997), or the general method of moments technique, for example CKLS (1992) and Tse (1995).

As Nowman pointed out, the discretized model (2) neglects errors introduced as a result of time aggregation. The discretized error arises because equation (1) is only shorthand notation for the stochastic differential equation (SDE),

$$\int_0^t dr(s) = \int_0^t [\alpha + \beta r(s)]\, ds + \int_0^t \sigma r^{\gamma}(s) dz(s), \tag{4}$$

which is the correct representation of the stochastic process. The more formal approach to discretizing equation (4) is to first solve the SDE for *r(t)* and then to discretize the solution. Thus, the discretization of the model (1) should read

$$\begin{aligned} r(t) &= \frac{\alpha}{\beta}[e^{\beta} - 1] + e^{\beta} r(t-1) + \varepsilon(t) \\ \varepsilon(t) &= \int_{t-1}^{t} e^{\beta(t-s)} \sigma r^{\gamma}(s)\, dz \end{aligned} \tag{5}$$

where the conditional mean and variance of the error term are approximated by

$$E_{t-1}[\varepsilon(t)] = 0 \quad \text{ and } E_{t-1}[\varepsilon^2(t)] = \frac{\sigma^2}{2\beta}[e^{2\beta} - 1]\ r(t-1)^{2\gamma}. \tag{6}$$

(See Bergstrom [1984], Nowman [1997], and the Appendix for details.) Note that equation (5) is the exact solution of the general model (1). Furthermore, the disparity between the discrete-time approximation, equations (2) and (3), and the above solution, equations (5) and (6), lessens as the mean-reversion parameter, β, tends to zero.

3. Standard estimation techniques

Two of the main techniques that are used in the literature to estimate one-factor interest rate models are the method of maximum likelihood and the general method of moments.

3.1 Method of maximum likelihood

The method of maximum likelihood is a parametric estimation technique. Under the assumption that the probability density function of the data has a particular parametric form, the method ascertains which parameter value would yield the greatest likelihood of obtaining the observed data. In other words, the method chooses the probability density function under which the observed data would have the highest likelihood of occurring. The likelihood function is simply the joint density function of the sample data. For computational convenience, the method focuses on the log of the likelihood function—parameter estimates obtained from the likelihood function and the log of the likelihood function are identical.

The standard approach is to assume that the model errors, $\varepsilon(t)$, are conditionally normal. In this case, the log likelihood function (LLF) for (2) or (5) is given by

$$\mathrm{LLF}(\rho) = -\frac{1}{2}\sum_{t=1}^{n}\left\{\log(2\pi h_t) + \frac{\varepsilon^2(t)}{h_t}\right\} \tag{7}$$

where $h_t = \mathrm{E}_{t-1}[\varepsilon^2(t)]$ is the conditional variance, n is the number of observations, and $\rho = \{\alpha, \beta, \sigma, \gamma\}$ is the vector of model parameters. The LLF estimates of the model parameters are then given by

$$\hat{\rho} = \underset{\rho}{\arg\max}\ \{\mathrm{LLF}(\rho)\}, \tag{8}$$

where $\hat{\rho}$ is the parameter vector that generates the largest value of the LLF.

3.2 General method of moments

The general method of moments (GMM) of Hansen (1982) is appealing in that no parametric assumptions need be made about the distribution of the errors, and the GMM errors are asymptotically consistent. GMM is closely related to the classical method of moments and instrumental variable estimation. The classical method of moments uses moment restrictions to estimate model parameters. These restrictions can be written as population moments whose expectation is zero when evaluated at the true parameter values. One of the key concepts behind GMM is that there is a set of moment conditions involving the parameter vector such that the expected value of these conditions at the true parameter vector is zero. In instrumental variable estimation, the key idea is to find a set of instruments that is correlated with the regressors but uncorrelated with the error terms. In other words, the instrument vector must be orthogonal to the errors. Instrumental variable estimation can be cast in a GMM framework where the momentum conditions are given by the requirement that the instrument vector be orthogonal to the errors. Consequently, the moment conditions are also referred to as orthogonality conditions.

More formally, the GMM estimation framework is as follows: Let $f(r(t);\rho)$ be a $q \times 1$ vector of disturbances that satisfy the following set of q orthogonality conditions—these conditions are usually restrictions on the moments of the errors in the model:

$$E[f(r(t);\rho)] = 0 . \tag{9}$$

Under standard technical conditions such that the law of large numbers holds (see Hansen [1982] and Hamilton [1994]), the sample average,

$$g_n(\rho) = \frac{1}{n}\sum_{t=1}^{n} f(r(t);\rho) , \tag{10}$$

is a good approximation for the orthogonality conditions (9) if n is large. Now let

$$J_n(\rho) = g_n'(\rho)\, W_n\, g_n(\rho) \tag{11}$$

where W_n is a positive definite $q \times q$ weighting matrix. The GMM estimate of the model parameters is given by

$$\hat{\rho}_n = \underset{\rho}{\arg\min}\ \{J_n(\rho)\} . \tag{12}$$

Under fairly general conditions $\hat{\rho}_n$ is a consistent estimator of the true parameter vector. Hansen (1982) showed that an efficient choice of W_n is given by $W_n = \Omega^{-1}$ where

$$\Omega = \lim_{n \to \infty} n\mathrm{E}[g_n(\rho_0)\ g_n(\rho_0)'] = \lim_{j \to \infty} \sum_{j=-\infty}^{+\infty} \mathrm{E}[\mathrm{f(r(t);}\rho_0)\ \mathrm{f'}(r(t-j);\rho_0)] \tag{13}$$

and ρ_0 is the true value of ρ. Thus, in order to implement GMM, an estimator for Ω is required. A standard approach is to replace the true autocovariances with sample autocovariances. A popular choice for the estimator of Ω is the Newey-West (1987) estimator,

$$\hat{\Omega}_n = \hat{\Phi}_0 + \sum_{j=1}^{s} \frac{s-j}{s} (\hat{\Phi}_j + \hat{\Phi}'_j) , \tag{14}$$

where

$$\hat{\Phi}_j = \frac{1}{n} \sum_{t=j+1}^{n} \mathrm{f(r(t);}\hat{\rho})\ \mathrm{f'}(r(t-j);\hat{\rho}) . \tag{15}$$

The Newey-West estimator is both consistent and positive definite. In addition to this estimator, several other estimators have also been proposed (see Ogaki [1992] and Hamilton [1994] for reviews).

Note that, if the number of parameters is equal to the number of orthogonality conditions, then the system is said to be just-identified and the estimate of ρ is independent of W_n. In this case, the parameter estimates are given by simply solving $g_n(\rho) = 0$.

In the case where the number of orthogonality conditions, q, is greater than the number of model parameters, l_ρ, diagnostic testing can be conducted using Hansen’s J-test (also known as the over-identification test). The restrictions implied by the model can be tested using Hansen’s J-test, which states that

$$nJ_n(\hat{\rho}) \overset{a}{\sim} \chi^2(q - l_\rho) . \tag{16}$$

Expanding upon the work of CKLS, choose the vector f($r(t)$;ρ) to be

$$f(r(t);\rho) = \begin{bmatrix} \varepsilon(t+1) \\ \varepsilon^2(t+1) - \frac{\sigma^2}{2\beta}[e^{2\beta} - 1] \ r(t)^{2\gamma} \end{bmatrix} \otimes Z_t \tag{17}$$

where $\varepsilon(t+1) = r(t+1) - e^{\beta} r(t) - \frac{\alpha}{\beta}[e^{\beta} - 1]$ and $Z_t = [1, r(t)]'$ is a vector of instruments, and $\otimes$ represents the Kronecker product.

4. The data and initial empirical results

The present study uses 90-day commercial paper rates for Canada and 3-month Treasury bill rates[2] for the United States. The data are recorded at a weekly frequency for the period 1975 to 1995 and consist of 1095 observations. Wednesday closing observations are used; Thursdays are used if the Wednesday was unavailable; and Tuesdays are used if both the Wednesday and the Thursday were unavailable.

Two main questions are addressed in this section: Is there a difference between the results obtained by estimating the discrete model, (2)-(3), and those obtained by estimating the aggregate model, (5)-(6)? Do the results depend on estimation technique?

Table 1 contains the results of estimating both the discrete model and the aggregate model via LLF and GMM methods. As shown in the table, there is virtually no difference between the estimates of the discrete and aggregate models for the Canadian 90-day commercial paper rate or for the 3-month U.S. Treasury bill rate series using either technique. Thus, it would appear that the results are not driven by aggregation effects. The results are quite similar as estimates of β—the rate of adjustment towards the long-run mean of the short-term interest rate—are small and the aggregate model reduces to the discrete model in the limit that β tends to zero. Thus, the higher-order aggregation effects are quite small and can be neglected from all practical purposes. Hence, all further discussion in this section will focus on the results of the discrete model.

For Canadian commercial paper rates, neither α nor β are significant, although the LLF and the GMM methods yield similar point estimates. In particular, the LLF and GMM methods yield estimates of the long-run mean, $-\alpha/\beta$, of 8.97 per cent and 8.89 per cent, respectively. Both these

2. Treasury bill rates are used, as opposed to eurodollar rates, to allow direct comparison of the results with earlier studies.

values are below the sample mean of the data series, which is 9.90 per cent. The methods disagree on the estimates of σ and γ. The LLF estimates are σ = 0.108 and γ = 0.441 while the GMM estimates are σ = 0.029 and γ = 0.998. Both σ and γ are significant for the LLF estimation. However, only γ is significant for the GMM estimation. The discrepancies in the estimates will be further investigated later in the paper.

Table 1: One-factor models: LLF and GMM estimates

	CP-CAN				TB-US			
	Discrete LLF	Aggregate LLF	Discrete GMM	Aggregate GMM	Discrete LLF	Aggregrate LLF	Discrete GMM	Aggregate GMM
α	0.0332 (0.0254)	0.0332 (0.0248)	0.0311 (0.0482)	0.0341 (0.0499)	0.0102 (0.0121)	0.0102 (0.0122)	0.0435 (0.0417)	0.0435 (0.0367)
β	-0.0037 (0.0027)	-0.0037 (0.0027)	-0.0035 (0.0053)	-0.0036 (0.0056)	-0.0012 (0.0024)	-0.0012 (0.0025)	-0.0061 (0.0069)	-0.0061 (0.0063)
σ	0.108* (0.012)	0.108* (0.001)	0.029 (0.021)	0.027 (0.020)	0.012* (0.001)	0.012* (0.001)	0.007* (0.003)	0.007* (0.003)
γ	0.441* (0.048)	0.442* (0.004)	0.998* (0.288)	1.019* (0.293)	1.532* (0.048)	1.532* (0.047)	1.761* (0.161)	1.761* (0.179)
$-\alpha/\beta$	8.97	8.97	8.89	9.47	8.50	8.50	6.61	6.61

Corrected standard error appear in parentheses. (*) indicates that the coefficient is statistically significant at the 5 per cent level.

For U.S. Treasury bills, neither α nor β are significant. Furthermore, the LLF and GMM methods yield different estimates of the long-run mean, -α/β, of 8.50 per cent and 6.61 per cent, resepectively. These values lie either side of the sample mean, which is 7.71 per cent. Both the estimates of σ and γ are significant. The LLF estimates are σ = 0.012 and γ = 1.532 while the GMM estimates are σ = 0.007 and γ = 1.761. The estimates of σ and γ for Treasury bills appear to be robust; they agree within standard error. Furthermore, the estimates are similar to previous empirical findings. For example, Tse (1995) found γ = 1.728 for 3-month U.S. money market rates using monthly data from August 1976 to May 1994; and Brenner, Harjes, and Kroner (1996) found γ = 1.559 for 13-week Treasury bill yields using weekly data from 9 February 1973 to 6 July 1990.

As a further robustness check on the different GMM estimates, different instrument sets were used. In addition to $Z_t = [1, r(t)]'$, the instruments sets $Z_t = [1, r(t), r(t-1)]'$ and

$Z_t = [1, r(t), r(t-1), r(t-2)]'$ were employed. The results for Canadian commercial paper and U.S. Treasury bills are reported in Tables 2 and 3, respectively. The estimates for Canadian commercial paper appear to be unstable, with estimates of γ ranging from to 0.229 to 0.998. Furthermore, Hansen's J-test rejects the model for two of the intrument sets. The estimates for U.S. Treasury bills are remarkably robust to the choice of instruments. Furthermore, the discrete model, (2)-(3), cannot be rejected using Hansen's J-test. A final observation is that it appears to be extremely difficult to estimate σ and γ independently. For the same data set, high estimates of γ are typically associated with lower estimates of σ. The difficulty of estimation may be a symptom of the conditional standard deviation being misspecified.

Table 2: One-factor models: Robustness to GMM instruments

	CP-CAN		
Z_t'	$[1, r_t]$	$[1, r_t, r_{t-1}]$	$[1, r_t, r_{t-1}, r_{t-2}]$
α	0.0311 (0.0482)	0.0533 (0.0448)	0.0331 (0.0417)
β	-0.0035 (0.0053)	-0.0064 (0.0048)	-0.0041 (0.0045)
σ	0.0290 (0.0206)	0.1260 (0.1123)	0.1497 (0.1176)
γ	0.9981* (0.2883)	0.3167 (0.3852)	0.2286 (0.3478)
$q - l_\rho$	0	2	4
$\chi^2(q - l_\rho)$	—	8.606**	11.770**

Corrected standard error appear in parentheses. (*) indicates that the coefficient is statistically significant at the 5 per cent level. (**) indicates rejection at the 5 per cent level. Standard errors are computed with a Newey-West estimator with s = 3. The number of degrees of freedom, $q - l_\rho$, is equal to the number of orthogonality conditions minus the number of parameters to be estimated. The χ^2 value refers to Hansen's J-test.

One of the drawbacks of the GMM technique is that a set of orthogonality conditions has to be specified. This entails both deciding on moment conditions and choosing a set of instruments. Thus, the rejection of the model for Canadian commercial paper could possibly be due to the fact that a poor set of instruments was chosen. For example, Kogure (1997) and Gouriéroux and Monfort (1996) considered the more general SDE,

$$dr(t) = a[r(t),\rho]dt + b[r(t),\rho]dz \ , \tag{18}$$

where ρ represents a vector of unknown parameters. They noted that moment conditions for (18) are given by

$$\mathrm{E}\left[h'[r(t)]\ \ a[r(t),\rho] + \frac{1}{2}h''[r(t)]\ \ b^2[r(t),\rho]\right] = 0 \ , \tag{19}$$

where $h: I\!R \to I\!R$ is any twice continuously differentiable function such that $\mathrm{E}[|h[r(t)]|] < \infty$. How does one go about deciding which orthogonality conditions to use from the above infinite set of moment conditions?

Table 3: One-factor Models: Robustness to GMM instruments

	TB-US		
Z_t'	$[1, r_t]$	$[1, r_t, r_{t-1}]$	$[1, r_t, r_{t-1}, r_{t-2}]$
α	0.0435 (0.0417)	0.0252 (0.0401)	0.0243 (0.0398)
β	-0.0061 (0.0069)	-0.0027 (0.0066)	-0.0027 (0.0065)
σ	0.0072* (0.0027)	0.0071* (0.0029)	0.0079* (0.0028)
γ	1.761* (0.161)	1.750 (0.172)	1.706 (0.156)
$q - l_\rho$	0	2	4
$\chi^2(q - l_\rho)$	—	2.633	3.019

Corrected standard error appear in parentheses. (*) indicates that the coefficient is statistically significant at the 5 per cent level. (**) indicates rejection at the 5 per cent level. Standard errors are computed with a Newey-West estimator with s = 3.

Gallant and Tauchen (1996) addressed the problem of which orthogonality conditions to choose in their paper entitled, “Which Moments to Match?” There, they formulated a systematic approach to generating moment conditions, which they called the efficient method of moments

(EMM). The remainder of the paper concentrates on using the EMM technique to further examine the validity of the one-factor model, (1), for Canadian commercial paper and U.S. Treasury bills.

5. Efficient method of moments (EMM)

The efficient method of moments specifically addresses the question of which set of moment conditions to match. A systematic approach to choosing the orthogonality conditions for the GMM estimator is given. The general idea is to first formulate a good statistical description of the data and then test if the structural model could possibly yield a similar description of the data; the structural model is the data-generating process that is postulated to have generated the observed data. EMM was initially proposed by Bansal, Gallant, Hussey, and Tauchen (1994; 1995), and further developed by Gallant and Tauchen (1996). Good overviews can be found in Tauchen (1995), and Gallant and Tauchen (1997a).

First, start with an auxiliary model that gives a good statistical description of the data. Suppose that the conditional density of r_t [$\equiv r(t)$] is given by

$$f(r_t|x_{t-1}, \theta) , \tag{20}$$

where $x_{t-1} = (r_{t-L}, \ldots, r_{t-2}, r_{t-1})$ and θ is a vector of parameters, of length l_θ, that characterizes the auxiliary model. The parameters θ are estimated by quasi-maximum likelihood using the scores of the auxiliary model (hence, the auxiliary model is also referred to as the score generator):

$$\hat{\theta}_n = \underset{\theta}{\arg\max} \frac{1}{n} \sum_{t=L+1}^{n} \log[f(r_t|x_{t-1}, \theta)] . \tag{21}$$

Now consider the structural model for r_t that depends on a vector of parameters ρ, of length l_ρ. Define the moment criterion,

$$m(\rho, \theta) = \mathrm{E}\left[\frac{\partial}{\partial\theta} \log[f(r_t(\rho)|x_{t-1}(\rho), \theta)]\right] , \tag{22}$$

where the expectation is with respect to the true density of the structural model. The expectation can be calculated by Monte Carlo methods, namely, averaging over a long simulation,

$$m(\rho, \hat{\theta}_n) = \frac{1}{N}\sum_{\tau=1}^{N}\frac{\partial}{\partial\theta}\log[f(\hat{r}_\tau(\rho)|\hat{x}_{\tau-1}(\rho), \hat{\theta}_n)] , \tag{23}$$

where $\{\hat{r}_t(\rho)\}_{\tau=1}^{N}$ is a simulation generated from the structural model. If the observed data is generated by the structural model, then one would expect that the moment criterion would be close to zero. More formally, the EMM estimator for the parameters of the structural model is given by

$$\hat{\rho}_n = \underset{\rho}{\arg\min}\ m(\rho, \hat{\theta}_n)'(\tilde{I}_n)^{-1}m(\rho, \hat{\theta}_n), \tag{24}$$

where

$$\tilde{I}_n = \frac{1}{n}\sum_{t=L+1}^{n}\left[\frac{\partial}{\partial\theta}\log[f(r_t|x_{t-1}, \hat{\theta}_n)]\right]\left[\frac{\partial}{\partial\theta}\log[f(r_t|x_{t-1}, \hat{\theta}_n)]\right]' . \tag{25}$$

The above choice of the weighting matrix is valid under the assumption that the auxiliary model yields a good statistical description of the data generator process, which is assumed to be the case. Note that the weighting matrix depends only on the observed data and is independent of the parameters of the structural model.

The validity of the structural model can be tested by noting that

$$n\,m(\rho, \hat{\theta}_n)'(\tilde{I}_n)^{-1}m(\rho, \hat{\theta}_n) \overset{a}{\sim} \chi^2(l_\theta - l_\rho) \tag{26}$$

under the null hypothesis that the structural model is correct. If the null hypothesis is rejected, then diagnostic testing can be carried out to see where the model fails. The *t*-statistics,

$$\hat{T}_n = (S_n)^{-1}\sqrt{n}m(\rho, \hat{\theta}_n) , \tag{27}$$

where

$$S_n = \left\{\text{diag}[\tilde{I}_n - \hat{M}_n[\hat{M}_n{}'(\tilde{I}_n)^{-1}\hat{M}_n]^{-1}\hat{M}_n{}']\right\}^{1/2} \text{ with } \hat{M}_n = \frac{\partial}{\partial\rho}m(\hat{\rho}_n, \hat{\theta}_n), \tag{28}$$

contain information about how well the structural model fits the scores of the auxiliary model. (A computationally less intensive diagnostic test is given by the quasi–*t*-statistic. The quasi–*t*-statistic is given by equation (27) with $S_n = \{\text{diag}[\tilde{I}_n]\}^{1/2}$. The quasi–*t*-statistic has a downward bias relative to 2 compared to the adjusted *t*-statistic.) Thus, large *t*-statistics, typically greater than 2, indicate those characteristics of the data that not satisfactorily explained by the structural model.

6. Semi–non-parametric (SNP) model

EMM estimation relies on choosing an auxiliary model that yields a good statistical description of the data. The conditional density of the interest rate process is estimated by the semi–non-parametric (SNP) approach. The SNP approach was developed by Gallant and Tauchen as a method for describing the properties of time series data. The conditional density of a multivariate process can be approximated by a Hermite polynomial series expansion around the standard normal density. The approach yields a non-linear, non-parametric model that nests several well-known models. For example, the model nests the Gaussian VAR model, the semi-parametric VAR model, the Gaussian ARCH model, and the semi-parametric ARCH model. In the SNP approach, the conditional density for the interest rate process, r_t, is

$$f(r_t | x_t, \theta) = \frac{[P(z_t, x_{t-1})]^2 \phi(z_t)}{\int [P(u, x_{t-1})]^2 \phi(u) \, du} \frac{1}{R_t} , \tag{29}$$

where $\phi(\bullet)$ is the standard normal density and

$$x_{t-1} = (r_{t-L}, \ldots, r_{t-2}, r_{t-1}) , \tag{30}$$

$$z_t = \frac{r_t - \mu_t}{R_t} , \tag{31}$$

$$\mu_t = \phi_0 + \sum_{j=1}^{L_u} \phi_j r_{t-j} , \tag{32}$$

$$R_t = \theta_0 + \sum_{j=1}^{L_r} \theta_j |z_{t-j}| , \tag{33}$$

$$P(z_t, x_{t-1}) = \sum_{\alpha=0}^{K_z} \left(\sum_{\beta=0}^{K_x} a_{\alpha\beta} x_{t-1}^{\beta} \right) z_t^{\alpha} \qquad \text{where} \qquad a_{00} = 1 , \tag{34}$$

$$\text{and } a_{\alpha\beta} x_{t-1}^{\beta} \equiv \sum_{\underline{\beta} \text{ s.t. } |\underline{\beta}| = \beta} a_{\alpha, \underline{\beta}} \prod_{j=1}^{L_p} r_{t-j}^{\beta_j} . \tag{35}$$

$\underline{\beta}$ is an L_p-vector, $\underline{\beta} = (\beta_1, \ldots, \beta_{L_p})$, such that β_j are non-negative integers and $|\underline{\beta}| = \sum_{j=1}^{L_p} \beta_j$.

The possible classes of auxiliary models can classified by the parmeters L_u , L_r , $L_{p,}$ K_z , and K_x. L_u is the number of lags in the linear part of the SNP model; L_r is the number of lags in the ARCH part; L_p is the number of lags of r_t that are included in the *x* part of *P(z,x);* K_z is the degree of the polynomial *P(z,x)* in *z;* and K_x is the degree of the polynomial *P(z,x)* in *x*. (By convention, $L_p = 1$ if $K_x = 0$.) The various SNP models can be classified according to whether the parameters L_u, L_r, L_p, K_z, and K_x are zero or non-zero. Some of the possible SNP models are outlined in Table 4.

Table 4: Classification of SNP models

SNP Model	L_u	L_r	L_p	K_z	K_x
Gaussian VAR	≥ 1	= 0	≥ 0	= 0	= 0
Semi-parametric VAR	≥ 1	= 0	≥ 0	≥ 1	= 0
Gaussian ARCH	≥ 0	≥ 1	≥ 0	= 0	= 0
Semi-parametric ARCH	≥ 0	≥ 1	≥ 0	≥ 1	= 0
Non-linear, non-parametric	≥ 0	≥ 0	≥ 1	≥ 1	≥ 1

Model selection is determined by examining the Schwarz Bayes (BIC), Hannan-Quinn (HQ), and Akaike (AIC) information. Thus the model that yields the lowest criteria values is selected. The BIC, HQ, and AIC information criteria are as follows:

$$\begin{aligned} \text{BIC} &= s_n(\hat{\theta}_n) + \frac{1}{2}(p_\theta / n) \log n \\ \text{HQ} &= s_n(\hat{\theta}_n) + (p_\theta / n) \log(\log n) \\ \text{AIC} &= s_n(\hat{\theta}_n) + p_\theta / n \end{aligned} \quad . \tag{36}$$

Each of the criteria has two components: the first is the value of the objective function that rewards for goodness of fit; the second is an adjustment that penalizes for too many parameters. The AIC tends to overparameterize the model while the BIC offers a sparser selection. The HQ criteria yield models between the BIC and AIC. Model selection is based on the BIC information criteria but the HQ and AIC values are listed for comparison. The values L_u, L_r, K_z, K_x, and L_p are determined sequentially. This approach is not necessarily optimal; however, it has the desirable property of sequentially refining the conditional density.

For the CP-Can series, the SNP model with $L_u = 2$, $L_r = 7$, $L_p = 1$, $K_z = 8$, and $K_x = 0$ is optimal under the BIC criteria. This model has 19 degrees of freedom. For the TB-US series, the SNP model with $L_u = 2$, $L_r = 9$, $L_p = 1$, $K_z = 4$, and $K_x = 0$ is optimal under the BIC criteria. This model has 17 degrees of freedom. Hence, semi-parametric ARCH models appear to be adequate to describe both sets of data.

The one-step-ahead conditional mean and conditional standard deviation for the above SNP fits are plotted in Figures 1 and 2, along with the raw data sets. The conditional means appear to track the series quite well and the conditional standard deviations also appear to track movements in the series.

Table 5: SNP estimation of the conditional density of CP-CAN

SNP tuning parameters						Objective function			
L_u	L_r	L_p	K_z	K_x	l_θ	s_n	BIC	HQ	AIC
1	0	1	0	0	3	-0.9732	-0.9635	-0.9678	-0.9704
2	0	1	0	0	4	-0.9822	-0.9822	-0.9879	-0.9914
3	0	1	0	0	5	-0.9982	-0.9820	-0.9892	-0.9935
4	0	1	0	0	6	-0.9982	-0.9820	-0.9892	-0.9935
2	1	1	0	0	5	-1.1300	-1.1138	-1.1210	-1.1254
2	2	1	0	0	6	-1.1954	-1.1760	-1.1846	-1.1898
2	3	1	0	0	7	-1.2258	-1.2032	-1.2132	-1.2193
2	4	1	0	0	8	-1.2383	-1.2124	-1.2239	-1.2309
2	5	1	0	0	9	-1.2477	-1.2186	-1.2315	-1.2394
2	6	1	0	0	10	-1.2489	-1.2165	-1.2309	-1.2396
2	7	1	0	0	11	-1.2557	-1.2201	-1.2359	-1.2455
2	8	1	0	0	12	-1.2558	-1.2170	-1.2342	-1.2447
2	7	1	4	0	15	-1.3792	-1.3307	-1.3522	-1.3653
2	7	1	5	0	16	-1.3787	-1.3270	-1.3496	-1.3639
2	7	1	6	0	17	-1.4014	-1.3464	-1.3708	-1.3856
2	7	1	7	0	18	-1.4022	-1.3440	-1.3698	-1.3855
2	7	1	8	0	19	-1.4109	-1.3495	-1.3767	-1.3933
2	7	1	9	0	20	-1.4101	-1.3455	-1.3741	-1.3916
2	7	1	8	1	25	-1.4220	-1.3315	-1.3716	-1.3961
2	7	1	8	2	34	-1.4270	-1.3075	-1.3605	-1.3928
2	7	2	8	1	34	-1.4296	-1.3101	-1.3630	-1.3954
2	7	3	8	1	43	-1.4334	-1.2848	-1.3507	-1.3909

L_u is the number of lags in the linear part of the SNP model; L_r is the number of lags in the ARCH part; L_p is the number of lags in the polynomial part, *P(z,x)*. The polynomial *P(z,x)* is of degree K_z in *z* and K_x in *x*; by convention, L_p =1 if K_x = 0. The number of SNP parameters is l_θ. The value of the objective function is s_n.The BIC, HQ, and AIC information criteria are listed.

Figure 1: Canadian commercial paper rates

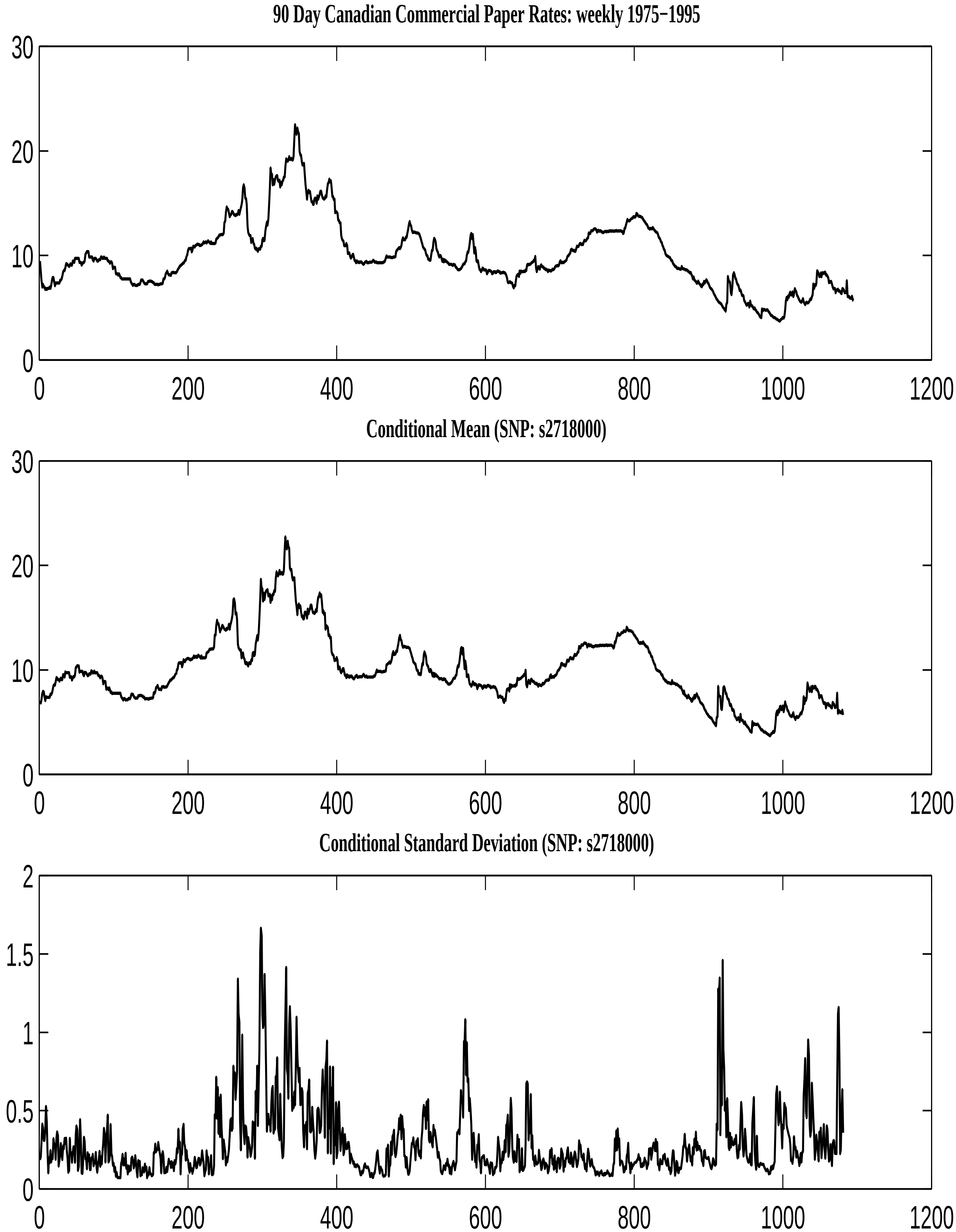
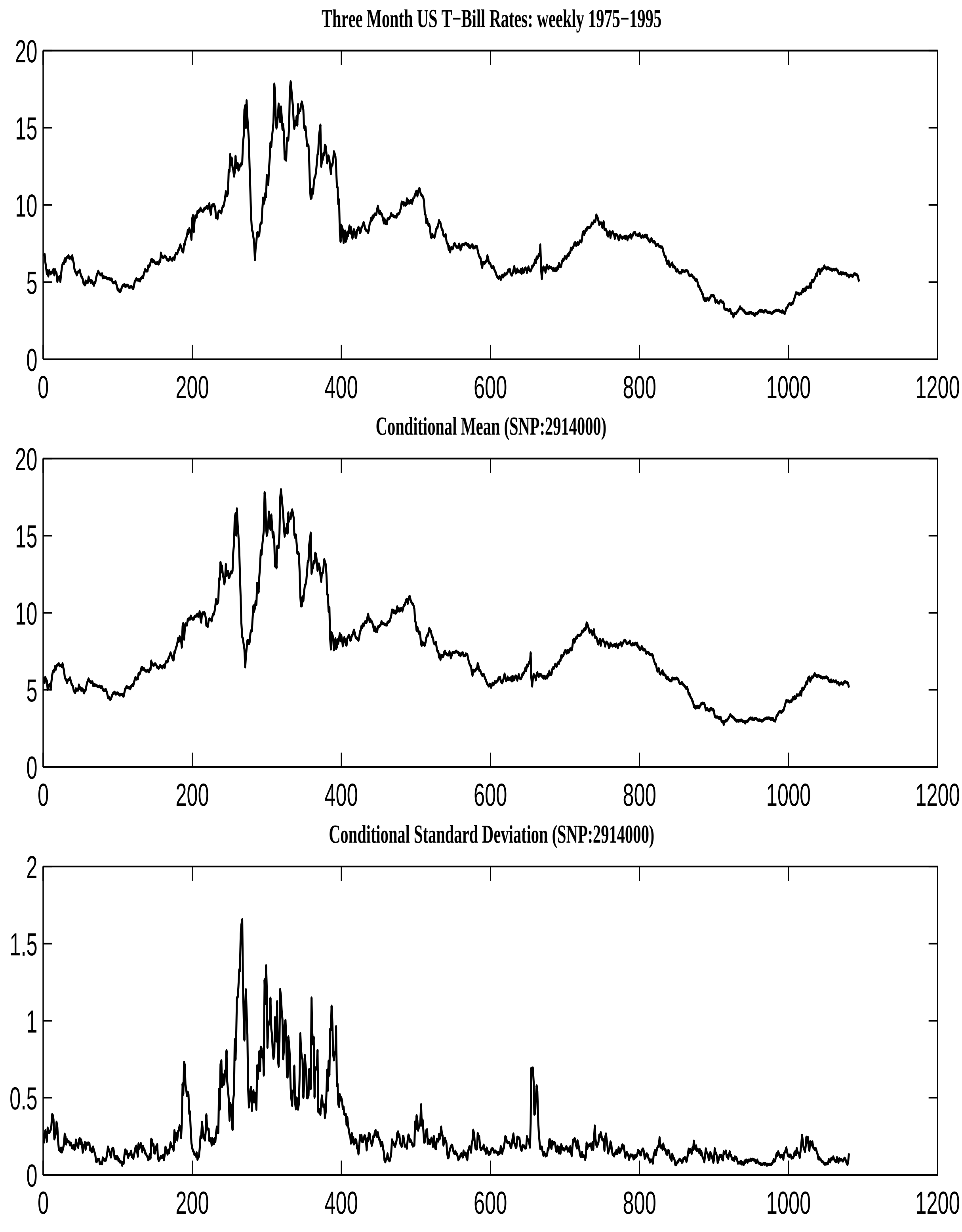


Table 6: SNP estimation of the conditional density of TB-US

SNP tuning parameters						Objective function			
L_u	L_r	L_p	K_z	K_x	l_θ	s_n	BIC	HQ	AIC
1	0	1	0	0	3	-0.7661	-0.7565	-0.7608	-0.7634
2	0	1	0	0	4	-0.7702	-0.7573	-0.7630	-0.7665
3	0	1	0	0	5	-0.7711	-0.7550	-0.7621	-0.7665
2	1	1	0	0	5	-1.0782	-1.0620	-1.0692	-1.0735
2	2	1	0	0	6	-1.1761	-1.1567	-1.1653	-1.1706
2	3	1	0	0	7	-1.2368	-1.2142	-1.2242	-1.2303
2	4	1	0	0	8	-1.2718	-1.2459	-1.2574	-1.2644
2	5	1	0	0	9	-1.2781	-1.2490	-1.2619	-1.2697
2	6	1	0	0	10	-1.2929	-1.2606	-1.2750	-1.2837
2	7	1	0	0	11	-1.2979	-1.2624	-1.2782	-1.2878
2	8	1	0	0	12	-1.3014	1.2627	-1.2798	-1.2903
2	9	1	0	0	13	-1.3069	-1.2649	-1.2835	-1.2949
2	10	1	0	0	14	-1.3069	-1.2617	-1.2817	-1.2940
2	9	1	4	0	17	-1.3556	-1.3007	-1.3250	-1.3399
2	9	1	5	0	18	-1.3562	-1.2981	-1.3239	-1.3396
2	9	1	4	1	19	-1.3645	-1.2934	-1.3249	-1.3441
2	9	1	4	2	24	-1.3706	-1.2834	-1.3221	-1.3457
2	9	2	4	1	27	-1.3700	-1.2828	-1.3215	-1.3451
2	9	3	4	1	32	-1.3700	-1.2697	-1.3156	-1.3435

L_u is the number of lags in the linear part of the SNP model; L_r is the number of lags in the ARCH part; L_p is the number of lags in the polynomial part, *P(z,x)*. The polynomial *P(z,x)* is of degree K_z in z and K_x in x; by convention, L_p =1 if K_x = 0. The number of SNP parameters is l_θ. The value of the objective function is s_n.The BIC, HQ, and AIC information criteria are listed.

Figure 2: U.S. Treasury bill rates



7. Efficient method of moments estimation

The first step in the EMM method is to estimate the auxiliary model. The second is to generate a simulated time series for the SDE. Given the results of Section 4, the SDE (1) was simulated using a simple discrete-time Euler scheme,[3]

$$r(t+h) = r(t) + [\alpha + \beta r(t)]h + \sigma\, r(t)^{\gamma} \sqrt{h}\ \ \varepsilon(t) \tag{37}$$

where $\varepsilon(t)$ is a random draw from the standard normal distribution. Each week was divided into 50 subintervals of equal length, h = 1/50. A simulation of weekly data of length 77,000 was generated. The first 2,000 generated observations were discarded so that the simulation used in the EMM estimation would be independent of the initial conditions. Thus, the simulation used in the EMM estimation had 75,000 observations at a weekly frequency. The simulation length was deemed to be long enough so that Monte Carlo errors become negligible.

For robustness, the estimation was conducted using different initial conditions, two different random number seeds for the simulation series, different numbers of weekly subintervals, and different simulation lengths. The results are fairly robust to the random number seeds, the number of subintervals, and the length of the simulation. However, the estimations were quite sensitive to the initial conditions. Robust results were found by using the results for the LLF and GMM estimations as initial conditions. Only the LLF and GMM estimates for α, β, and σ were used. In addition, a wide range of initial values for γ was also used to prevent biasing the EMM estimates of γ towards the LLF and GMM estimates of γ.

For Canadian 90-day commercial paper rates, none of the structural model parameter EMM estimates are statistically significant (Table 7). The point estimates of α = 0.03143 and β = –0.007615 are smaller than the corresponding LLF and GMM estimates. The other EMM estimates are σ = 0.0000255 and γ = 2.6459. The γ estimate is much larger than any of the previous estimates. Not surprisingly, the χ^2-test overwhelmingly rejects the one-factor model, (1), as offering an adequate description of the evolution of 90-day Canadian commercial rates. The score generator diagnostics listed in Table 8 pinpoint the problem with the model. The one-factor model is unable to fit the scores associated with the Hermite polynomial coefficients (of the auxiliary model). In particular, the *t*-statistics of all the even-power Hermite polynomial coefficients are all significantly greater than 2. Thus, the model is unable to explain the non-Gaussian deviations of the data.

3. A weak order two scheme (see Kloeden and Platen [1992]) for the SDE was also implemented, yielding similar results. However, this method increased the estimation time by a factor of six.

Table 7: EMM estimates

	CP-CAN	TB-US
Score generator	s2718000	s2914000
α	0.00807 (0.00724)	0.0314* (0.0150)
β	-0.000710 (0.000673)	-0.00762 (0.00570)
σ	0.0000255 (0.0000881)	0.00280 (0.00247)
γ	2.646 (1.441)	2.425* (0.185)
$l_\theta - l_\rho$	15	13
$\chi^2(l_\theta - l_\rho)$	78.481	15.472

Corrected standard error appear in parentheses. (*) indicates that the coefficient is statistically significant at the 5 per cent level. The number of degrees of freedom, $l_\theta - l_\rho$, is equal to the number of auxiliary model parameters minus the number of parameters to be estimated. Some comparison χ^2 probabilities that are helpful in interpreting the above χ^2 tests are as follows: $\mathrm{Prob}[\chi^2(15) \geq 37.7] = 0.001$, $\mathrm{Prob}[\chi^2(13) \geq 16.0] = 0.25$ and $\mathrm{Prob}[\chi^2(13) \geq 19.8] = 0.10$.

The one-factor model does a much better job of modelling the evolution of the 3-month U.S. Treasury bill rates. The EMM estimates are $\alpha = 0.0314$, $\beta = -0.00762$, $\sigma = 0.00280$, and $\gamma = 2.425$. Only the estimates of α and γ are statistically significant. Furthermore, the EMM estimate of γ agrees with the GMM estimate within error bounds. However, the insignificance of the σ estimate casts some doubt on the validity of the γ estimate. The χ^2 -test is unable to reject the null hypothesis that the structural model, (1), gives a valid description of 3-month U.S. Treasury bill rates. However, the score generator diagnostics listed in Table 9 indicate that the one-factor model is unable to fit the score associated with the quartic term of the Hermite polynomial. Overall, it would appear that the one-factor model, (1), is better able to describe the 3-month U.S. Treasury bill rate series than the Canadian 90-day commercial paper series, althought the fit is far from perfect.

Table 8: CP-CAN score generator diagnostics

Score generator eqn.		Quasi-*t*-stat	*T*-stat
ϕ_0	Mean	-0.405	-1.226
ϕ_1	VAR	0.801	1.063
ϕ_2		0.939	1.227
θ_0	Var	0.939	1.840
θ_1	ARCH	0.370	0.413
θ_2		-0.996	-1.167
θ_3		-0.809	-0.909
θ_4		-1.102	-1.330
θ_5		0.948	1.088
θ_6		0.43	0.451
θ_7		0.592	0.643
a_{10}	Hermite	0.864	1.149
a_{20}		3.040	5.842
a_{30}		1.612	1.794
a_{40}		4.890	6.006
a_{50}		0.955	0.991
a_{60}		4.959	5.329
a_{70}		0.685	0.695
a_{80}		4.595	4.740

See equations (28)-(29) for a description of the calculation of the above *t*-statistics.

Table 9: TB-US Score generator diagnostics

Score generator eqn.	Quasi-*t*-stat	*T*-stat
ϕ_0 Mean	0.724	0.779
ϕ_1 VAR	-0.391	-1.132
ϕ_2	-0.408	-1.122
θ_0 Var	0.655	1.379
θ_1 ARCH	0.842	1.357
θ_2	0.830	1.126
θ_3	0.799	1.004
θ_4	-0.228	-0.334
θ_5	-0.590	-0.845
θ_6	0.621	0.736
θ_7	0.462	0.534
θ_8	0.305	0.438
θ_9	0.553	0.809
a_{10} Hermite	-0.368	-0.380
a_{20}	0.917	1.714
a_{30}	-0.903	-0.928
a_{40}	1.955	2.721

See equations (28)-(29) for a description of the calculation of the above *t*-statistics.

8. Conclusion and discussion

A broad class of one-factor models for the short-term interest rate in Canada and the United States was examined. The model parameter estimates were found to be independent of whether a discretized version or a more accurate aggregation version of the continuous-time model was used. Furthermore, the parameter estimates for the U.S. Treasury bill interest series are more or less independent of the estimation technique employed—the study compares results from the LLF, GMM, and EMM estimation methods. However, the results for the Canadian commercial paper rates series vary with the estimation technique. Hypothesis testing suggests that the class of one-factor models for the short-term interest rate, (1), are inadequate as models for the evolution of the Canadian commercial paper rate. The evidence against the model for U.S. Treasury bill rates is only marginal. Thus, the one-factor model, (1), appears to offer a much better description of U.S. data than similar Canadian data. This result casts doubt on the usefulness of one-factor models, such as the CIR square root model, for providing an adequate basis for analyzing the risks of Canadian-dollar fixed-income portfolios.

It is conceivable that the inconsistent results for Canada stem from the possibility that the estimation techniques (LLF, GMM, and EMM) are unable to correctly identify time series data that adhere to the one-factor model, (1). However, the analysis presented in Appendix B indicates that the difference is probably not due to the estimation techniques. Monte Carlo experiments suggest that the various estimation techniques are all able to correctly identify a time series whose data-generating process is truly given by equation (1). Thus, the discrepancies in the estimations are not a symptom of the estimation techniques but rather an indication that the class of one-factor interest rate models given by (1) is inadequate to describe the observed interest rate series.

Several authors have tried to patch up the class of one-factor models, (1), by replacing the parameter σ by a time-varying parameter, h_t (for examples, see Brenner, Harjes, and Kroner [1996], Andersen and Lund [1997], and Brailsford and Mahesweran [1998]). For example, a GARCH(1,1) specification for the time varying volatility parameter could be achieved by replacing equation (3) by $E_t[\varepsilon_{t+1}^2] = \sigma^2_{t+1} = h^2_{t+1} r_t^{2\gamma}$ where $h^2_{t+1} = \omega + \phi h^2_t + \delta \varepsilon_t^2$. Note that modelling volatility as a GARCH process allows volatility to be time varying but volatility is not modelled as separate random process. Thus, the GARCH extension would still be classified as a one-factor model. However, preliminary research indicates that even these models do not provide an adequate characterization of the process driving short-term interest rates in Canada.

Principal component analysis suggests that three factors are necessary to describe movements in the yield curve (Litterman, Scheinkman, and Weiss [1991] and Han [1997]). These factors are commonly referred to as shift, tilt, and twist factors and are associated with the level, slope, and curvature of the yield curve, respectively. Han (1997) found that these three factors were able to explain 94 per cent of the movements in the Canadian yield curve. Hence, in addition to the short-term interest rate, at least two other factors are necessary. One possibility for an appropriate second factor is stochastic volatility. Another possibility is a factor related to a time-varying long-run mean of short-term interest rates, which may more adequately capture mean-reversion of short-term interest rates.

Affine models of the term-structure of interest rates incorporating stochastic volatility, and a time-varying long-run mean for the short-term interest rates (which is a three-factor model of the term structure of interest rates) have been discussed by Dai and Singleton (1997). Balduzzi, Das, Foresi and Sundaram (1996), Chen (1996) and Gong and Remolona (1996) also discuss affine three-factor models for the term structure of interest rates. The authors intend to employ EMM techniques to investigate whether Chen's three-factor model offers a good description of the evolution of the Canadian term structure.

Appendix A: Solving the stochastic differential equation

Consider the stochastic differential equation

$$dr(t) = [\alpha + \beta\, r(t)]\, dt + \phi(r,t)\ dz\,. \tag{38}$$

The above equation can be solved by first introducing the variable $Y(t) = \alpha + \beta\, r(t)$ and then using Ito's lemma to notice that the following equation must hold

$$d[e^{-\beta t} Y(t)] = \beta e^{-\beta t} \phi(r,t)\ dz\ . \tag{39}$$

Equation (39) is simply shorthand notation for

$$[e^{-\beta s} Y(s)]\Big|_0^t = \int_0^t \beta e^{-\beta s} \phi(r,s) dz\,. \tag{40}$$

Simplifying equation (40) yields the general solution

$$r(t) = -\frac{\alpha}{\beta} + \frac{1}{\beta} e^{\beta t}\ [\alpha + \beta\, r(0)] + e^{\beta t} \int_0^t \beta e^{-\beta s} \phi(r,s) dz\,. \tag{41}$$

The solution can also be written in an iterative form, namely:

$$\begin{aligned} r(t) &= \frac{\alpha}{\beta}[\, e^{\beta} - 1] + e^{\beta}\ r(t-1) + \varepsilon(t) \\ \varepsilon(t) &= \int_{t-1}^{t} e^{\beta(t-s)} \phi(r,s) dz \end{aligned}\quad . \tag{42}$$

Thus, equation (42) is the correct discretization model for the stochastic differential equation (38). Equation (42) explicitly takes care of the aggregation over time issues. Note that the conditional mean and variance of the error term are given by

$$E_{t-1}[\varepsilon(t)] = 0 \tag{43}$$

and

$$E_{t-1}[\varepsilon^2(t)] = E_{t-1}\left[\left\{\int_{t-1}^{t} e^{\beta(t-s)}\phi(r,s)dz\right\}^2\right]$$

$$= E_{t-1}\left[\int_{t-1}^{t} e^{2\beta(t-s)}\phi^2(r,s)ds\right] \qquad \text{By Ito Isometry}. \tag{44}$$

$$= \int_{t-1}^{t} e^{2\beta(t-s)} E_{t-1}[\phi^2(r,s)]\, ds$$

To proceed further, let $\phi(r,t) = \sigma\, r(t)^{\gamma}$ and approximate $E_{t-1}[\phi^2(r,s)]$ by $\sigma^2\, r(t-1)^{2\gamma}$. The conditional variance of the error term can then be approximated by

$$E_{t-1}[\varepsilon^2(t)] = \frac{\sigma^2}{2\beta}[\, e^{2\beta} - 1]\;\; r(t-1)^{2\gamma}. \tag{45}$$

Appendix B: Monte Carlo experiment

A Monte Carlo experiment was conducted in order to gain some insight into the discrepancies between the LLF, GMM, and EMM estimation techniques that were noted in Sections 4 and 7. A weekly frequency interest series was generated using the SDE (1), with known parameter values, as the data-generating process. A simple discrete-time Euler scheme, with each week divided into 50 equally spaced subintervals, was employed to generate a simulation of a short-term interest series of weekly observations and total length 1095 observations. The following model parameters were chosen: $\alpha = 0.08$, $\beta = -0.01$, $\sigma = 0.01$, and $\gamma = 1.5$. Hence, the generated series should have a long-run mean of 8 per cent and a conditional volatility that is dependent on the level of the series.

The simulated time series is graphed in the upper panel of Figure 3. The middle and lower panels of Figure 3 show the one-step-ahead condition mean and conditional standard deviation. The lower panel shows a clear dependence of the conditional volatility on the level.

The results of the LLF, GMM, and EMM estimations[4] are presented in Table 10. All three methods give similar, statistically significant estimates of σ and γ. The GMM method yields the best estimates of σ and γ in terms of both closeness to the true parameter values and smallness of the standard deviations of the estimates. In terms of mean absolute percentage error (MAPE)[5], the relative ranking of the three estimation techniques is: first, LLF; second, EMM; and third, GMM. (It is not too surprising that the LLF method wins out since the data were generated using normally distributed errors.) However, the LLF estimates of α and β have relatively large standard deviations, which are comparable to the EMM standard deviations. The GMM method overestimates the values of α and β. The GMM estimates are roughly twice the size of the true values. The EMM estimates of α and β are statistically indistinguishable from the true values.

4. The GMM estimate presented is for the just-identified system with the instrumental vector $Z_t = [1, r_t]'$ (see Section 3.2). The EMM estimate uses a simulation of weekly data of length 77,000 observations of which the first 2,000 observations are discarded. As before, each week is divided into 50 subintervals. A SNP auxiliary model with $L_u = 1$, $L_r = 5$, $L_p = 1$, $K_z=5$, and $K_x = 0$ best describes the statistical properties of the Monte Carlo time series.

5. MAPE = $\frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{\rho}_i - \rho_i}{\rho_i}\right|$ where $\hat{\rho}_i$ are the parameter estimates, ρ_i are the true parameter values, and N is the number of parameters.

Furthermore, the EMM χ^2-test is unable to reject the structural model (1), and the score generator diagnostics (see Table 11) only offer extremely marginal evidence against the model.[6]

The above results suggest that all three estimation techniques, LLF, GMM and EMM, are able to identify the one-factor model (1) when it is the true data-generating process. Of course, this result is only indicative since it is based on a single Monte Carlo trial.

Table 10: Simulation results

	Simulation		
	LLF	GMM	EMM
α	0.0825* (0.0385)	0.1592* (0.0140)	0.1176* (0.0310)
β	-0.0084 (0.0046)	-0.0173* (0.0014)	-0.0134* (0.0039)
σ	0.00926* (0.00161)	0.00998* (0.00126)	0.00824* (0.00149)
γ	1.536* (0.077)	1.504* (0.048)	1.575* (0.091)
$-\alpha/\beta$	9.82	9.20	8.78
$\chi^2(15)$	-	-	15.978

Corrected standard error appear in parentheses. (*) indicates that the coefficient is statistically significant at the 5 per cent level. The simulated data was generated using the data generating process (1) with parameter values $\alpha = 0.08$, $\beta = -0.01$, $\sigma = 0.01$ and $\gamma = 1.5$. Some comparison χ^2 probabilities that are helpful in interpreting the above χ^2 tests are as follows: $\text{Prob}[\chi^2(15) \geq 14.3] = 0.50$, $\text{Prob}[\chi^2(15) \geq 18.2] = 0.25$ and $\text{Prob}[\chi^2(15) \geq 25.0] = 0.05$.

6. More weight is given to the χ^2-test since it is a joint hypothesis test for all the score generator equations. The t-statistic tests are singular hypothesis tests that do not use the full information set, i.e., the information in the off-diagonal elements of $\hat{S}_n$ is not used in the t-statistic test (see Section 5).

Figure 3: Simulated data: Input to SNP

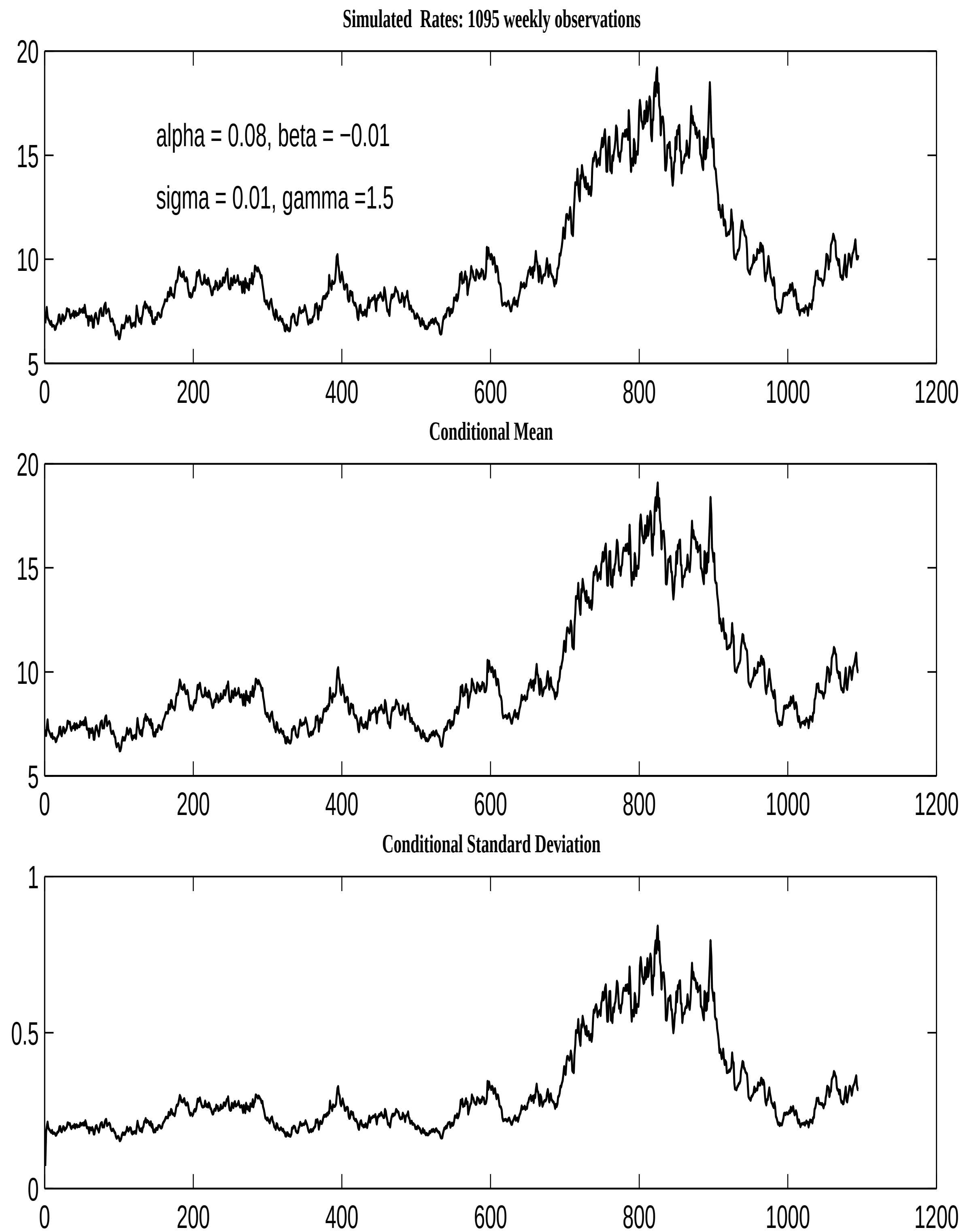


Figure 4: SNP Output: s1514010.fit

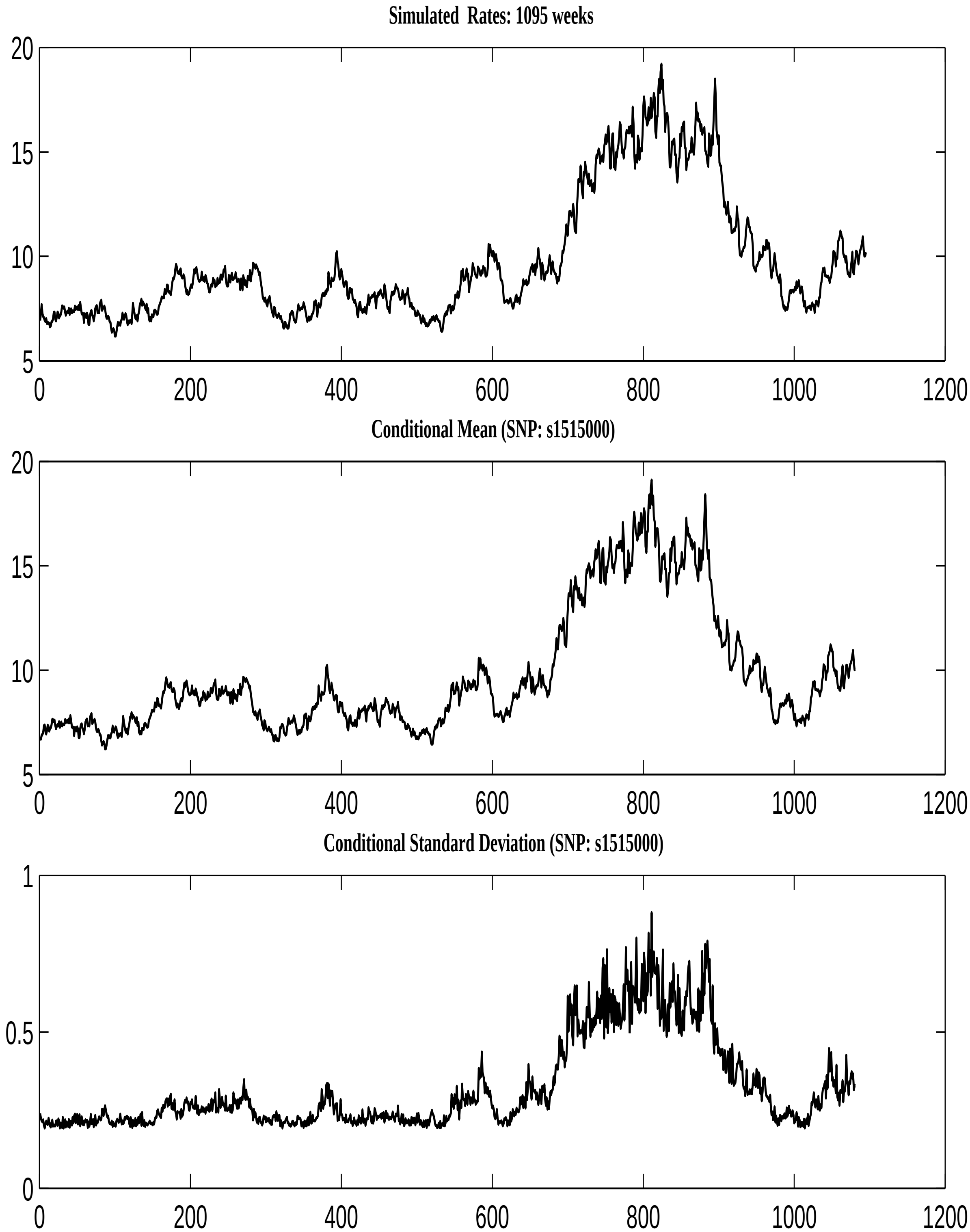


Table 11: Simulation Score generator diagnostics

Score generator eqn.	Quasi-*t*-stat	*T*-stat
ϕ_0 Mean	0.981	1.017
ϕ_1 VAR	-0.414	-0.419
θ_0 Var	0.353	0.495
θ_1 ARCH	1.476	1.610
θ_2	0.447	0.532
θ_3	0.401	0.511
θ_4	0.509	0.544
θ_5	-0.112	-0.126
a_{10} Hermite	-0.948	-1.031
a_{20}	1.198	1.312
a_{30}	-0.223	-0.226
a_{40}	0.293	0.433
a_{50}	-0.331	-0.362
a_{60}	1.847	2.025
a_{70}	0.592	0.631
a_{80}	0.955	1.536
a_{90}	0.724	0.824
$a_{10,0}$	1.564	1.738
$a_{11,0}$	0.613	0.641

See equations (28)-(29) for a description of the calculation of the above *t*-statistics.

Bibliography

Andersen, T. G. and J. Lund. 1997. "Estimating continuous-time stochastic volatility models of the short-term interest rate." *Journal of Econometrics* 77: 343–377.

Balduzzi, P., S. R. Das, S. Foresi, and R. Sundaram. 1996. "A Simple Approach to Three-Factor Affine Models of the Term Structure." *Journal of Fixed Income* 6: 43–53.

Bansal, R., A. R. Gallant, R. Hussey, and G. Tauchen. 1994. "Computational Aspects of Nonparametric Simulation Estimation." In *Computational Techniques for Econometrics and Economic Analysis,* edited by D. A. Belsley, 3–22. Boston: Kluwer Academic.

———. 1995. "Nonparametric estimation of structural models for high-frequency currency market data." *Journal of Econometrics* 66: 251–287.

Bergstrom, A. R. 1984. "Continuous Time Stochastic Models and Issues of Aggregation over Time." *Handbook of Econometrics,* Vol. II, edited by Z. Friliches and M. D. Intriligator. Amsterdam: Elsevier Science.

Black, F. and B. Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81: 637–657.

Bliss, R. B. and D. C. Smith. 1997. "The Elasticity of Interest Rate Volatility: Chan, Karolyi, Longstaff, and Sanders Revisited." Federal Reserve Bank of Atlanta Working Paper 97-13a. [Formerly titled "The Stability of Interest Rate Processes."]

Brailsford, T. J. and K. Maheswaran. 1998. "The Dynamics of the Australian Short-Term Interest Rate." SSRN Working Paper 98062903.

Brennan, M. J. M. and S. Schwartz. 1980. "Analyzing Convertible Bonds." *Journal of Financial and Quantitative Analysis* 15: 907–929.

Brenner, R. J., R. H. Harjes, and K. F. Kroner. 1996. "Another Look at Models of Short-Term Interest Rate." *Journal of Financial and Quantitative Analysis* 31: 85–107.

Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders. 1992. "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate." *Journal of Finance* 47: 1209–1227.

Chen, L. 1996. "Stochastic Mean and Stochastic Volatility—A Three-Factor Model of the Term Structure of Interest Rates and Its Applications in Derivatives Pricing and Risk Management." *Financial Markets, Institutions & Instruments* 5(1): 1–88.

Cox, J. C. 1975. "Notes on Option Pricing I:Constant Elasticity of Variance Diffusion." Working paper, Stanford University.

Cox, J. C., J. E. Ingersoll, and S. A. Ross. 1980. "An Analysis of Variable Rate Loan Contracts." *Journal of Finance* 35: 389–403.

———. 1985. "A Theory of the Term Structure of Interest Rates." *Econometrica* 53: 385–407.

Dahlquist, M. 1996. "On Alternative Interest Rate Processes." *Journal of Banking and Finance* 20: 1093–1119.

Dai, Q. and K. J. Singleton. 1997. "Specification Analysis of Affine Term Structure Models." NBER Working Paper 6128.

Dothan, U. L. 1978. “On the Term Structure of Interest Rates.” *Journal of Financial Economics* 6: 59–69.

Duguay, P. 1994. “Empirical Evidence on the Strength of the Monetary Transmission Mechanism in Canada: An Aggregate Approach.” *Journal of Monetary Economics* 33: 39–61.

Gallant, A. R. and G. Tauchen. 1996. “Which Moments to Match?” *Econometric Theory* 12: 657–681.

———. 1997a. “EMM: A Program for Efficient Method of Moments Estimation.” Available along with code via anonymous ftp at ftp.econ.duke.edu in the subdirectory home/get/emm.

———. 1997b. “SNP: A Program for Nonparametric Time Series Analysis” Available along with code via anonymous ftp at ftp.econ.duke.edu in the subdirectory home/arg/snp.

Gong, F. F. and E. M. Remolona. 1996. “A Three-Factor Econometric Model of the U.S. Term Structure.” Federal Reserve Bank of New York Research Paper No. 9619.

Gouriéroux, C. and A. Monfort. 1996. *Simulation-Based Econometric Methods.* Oxford: Oxford University Press.

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

Han, B. 1997. “A Principal Components Analysis of the Canadian Yield Curve.” Bank of Canada, mimeograph.

Hansen, L. P. 1982. “Large Sample Properties of Generalized Method of Moments Estimators.” *Econometrica* 50: 1029–1054.

Kloeden, P. E. and E. Platen. 1992. *Numerical Solution of Stochastic Differential Equations.* Vol. 23 of *Applications of Mathematics*, edited by A. V. Balakrishnan, I. Karatzas, and M. Yor. New York: Springer-Verlag.

Kogure, A. 1997. “A New Approach to the Estimation of Stochastic Differential Equations with an Application to the Japanese Interest Rates.” Institute for Monetary and Economic Studies (IMES) Discussion Paper 97-E-8, Bank of Japan.

Litterman, R., J. Scheinkman, and L. Weiss. 1991. “Volatility and the Yield Curve” *Journal of Fixed Income* 1: 49–53.

Merton, R. C. 1973. “Theory of Rational Option Pricing.” *Bell Journal of Economics and Management Science* 4: 141–183.

Nowman, K. B. 1997. “Gaussian Estimation of Single-Factor Continuous Time Models of the Term Structure of Interest Rates.” *Journal of Finance* 52: 1695–1706.

Ogaki, M. 1992. “An Introduction to the Generalized Method of Moments.” University of Rochester Center for Economic Research Working Paper No. 314.

Tauchen, G. 1995. “New Minimum Chi-Square Methods in Empirical Finance.” Presented at the Seventh World Congress of the Econometric Society, Tokyo, 23–28 August.

Tse, Y. K. 1995. “Some International Evidence on Stochastic Behavior of Interest Rates.” *Journal of International Money and Finance* 14: 721–738.

Vasicek, O. 1977. “An Equilibrium Characterization of the Term Structure.” *Journal of Financial Economics* 5: 177–188.

Bank of Canada Working Papers
Documents de travail de la Banque du Canada

Working papers are generally published in the language of the author, with an abstract in both official languages. ***Les documents de travail sont publiés généralement dans la langue utilisée par les auteurs; ils sont cependant précédés d'un résumé bilingue.***

1999

99-18	Estimating One-Factor Models of Short-Term Interest Rates	D. Mc Manus and D. Watt
99-17	Canada's Exchange Rate Regime and North American Economic Integration: The Role of Risk-Sharing Mechanisms	Z. Antia, R. Djoudad, and P. St-Amant
99-16	Optimal Currency Areas: A Review of the Recent Literature	R. Lafrance and P. St-Amant
99-15	The Information Content of Interest Rate Futures Options	D. Mc Manus
99-14	The U.S. Capacity Utilization Rate: A New Estimation Approach	R. Lalonde
99-13	Indicator Models of Core Inflation for Canada	R. Dion
99-12	Why Canada Needs a Flexible Exchange Rate	J. Murray
99-11	Liquidity of the Government of Canada Securities Market: Stylized Facts and Some Market Microstructure Comparisons to the United States	T. Gravelle
99-10	Real Effects of Collapsing Exchange Rate Regimes: An Application to Mexico	P. Osakwe and L. Schembri
99-9	Measuring Potential Output within a State-Space Framework	M. Kichian
99-8	Monetary Rules When Economic Behaviour Changes	R. Amano, D. Coletti, and T. Macklem
99-7	The Exchange Rate Regime and Canada's Monetary Order	D. Laidler
99-6	Uncovering Inflation Expectations and Risk Premiums from Internationally Integrated Financial Markets	B.S.C. Fung, S. Mitnick, and E. Remolona
99-5	The Quantity of Money and Monetary Policy	D. Laidler
99-4	An Intraday Analysis of the Effectiveness of Foreign Exchange Intervention	N. Beattie and J-F. Fillion
99-3	Forecasting GDP Growth Using Artificial Neural Networks	G. Tkacz and S. Hu
99-2	Capital Gains and Inflation Taxes in a Life-cycle Model	C. Leung and G.-J. Zhang
99-1	Dynamic Employment and Hours Effects of Government Spending Shocks	M. Yuan and W. Li

Copies and a complete list of working papers are available from:
Pour obtenir des exemplaires et une liste complète des documents de travail, prière de s'adresser à:

Publications Distribution, Bank of Canada
234 Wellington Street Ottawa, Ontario K1A 0G9

Diffusion des publications, Banque du Canada
234, rue Wellington, Ottawa (Ontario) K1A 0G9

E-mail / Adresse électronique: publications@bank-banque-canada.ca
WWW: http://www.bank-banque-canada.ca/