

# Nonparametric Estimation and Specification Testing of a Two-Factor Interest Rate Model

**Brennan S. Thompson**

Ryerson University

[digital.library.ryerson.ca/object/121](https://digital.library.ryerson.ca/object/121)

Please Cite:

Thompson, B. S. (2009). Nonparametric estimation and specification testing of a two-factor interest rate model. *Economics Bulletin*, 29(3), 2343-2349.

## Volume 29, Issue 3

### Nonparametric estimation and specification testing of a two-factor interest rate model

Brennan S. Thompson  
*Ryerson University*

#### Abstract

We propose a simple, flexible approach to nonparametric estimation and specification testing for a two-factor interest rate model. These methods are illustrated with a Monte Carlo experiment and an empirical example.

---

The author wishes to Thanasis Stengos, Alex Maynard, and Yiguo Sun for their helpful comments.

**Citation:** Brennan S. Thompson, (2009) "Nonparametric estimation and specification testing of a two-factor interest rate model", *Economics Bulletin*, Vol. 29 no.3 pp. 2343-2349.

**Submitted:** Mar 05 2009. **Published:** September 16, 2009.

## 1. Introduction

Letting  $r_t$  be the short rate,  $l_t$  be the consol rate, and  $s_t = r_t - l_t$  be the spread, the two-factor continuous-time model of Schaefer and Schwartz (1984) is

$$ds_t = \mu_1(s_t, l_t)dt + \sigma_1(s_t, l_t)dW_{1,t} \quad (1)$$

$$dl_t = \mu_2(s_t, l_t)dt + \sigma_2(s_t, l_t)dW_{2,t}, \quad (2)$$

where  $\mu_k(\cdot)$  and  $\sigma_k(\cdot)$ , for  $k = 1, 2$ , are unknown drift and diffusion functions, respectively, and  $W_{1,t}$  and  $W_{2,t}$  are independent standard Wiener processes.

Knight et al. (2006) have proposed a semiparametric estimation method for this model, which requires placing linear identifying restrictions on  $\mu_1(\cdot)$ . In this note, we propose a flexible, fully nonparametric estimation method, as well as a procedure for testing parametric specifications of this model. Aside from not placing any restrictive assumptions on the model, our method has the advantage of utilizing local linear (LL) regression, which avoids the well-known boundary bias of local constant (Nadaraya-Watson) regression (Chapman and Pearson, 2000; Fan and Zhang, 2003).

## 2. Estimation

Suppose we have observations  $\{(s_{\Delta}, l_{\Delta}), (s_{2\Delta}, l_{2\Delta}), \dots, (s_{n\Delta}, l_{n\Delta})\}$  at dates  $\{t_1 = \Delta, t_2 = 2\Delta, \dots, t_n = n\Delta\}$  on the time interval  $[0, T]$ , where  $\Delta = T/n$  is the discretization step. Let  $Y_{1,i} = (s_{(i+1)\Delta} - s_{i\Delta})$ ,  $Y_{2,i} = (l_{(i+1)\Delta} - l_{i\Delta})$ ,  $X_{1,i} = s_{i\Delta}$ , and  $X_{2,i} = l_{i\Delta}$ , for  $i = 1, \dots, (n-1)$ . Using a Euler-type approximation, we can rewrite system (1) – (2) as

$$Y_{k,i} \approx \mu_k(X_{1,i}, X_{2,i})\Delta + \sigma_k(X_{1,i}, X_{2,i})\epsilon_{k,i}\sqrt{\Delta}, \quad k = 1, 2,$$

with  $E(\epsilon_{k,i}|X_{1,i}, X_{2,i}) = 0$ ,  $\text{Var}(\epsilon_{k,i}|X_{1,i}, X_{2,i}) = 1$ , and  $E(\epsilon_{1,i}\epsilon_{2,i}|X_{1,i}, X_{2,i}) = 0$ .

Since the errors in these equations are uncorrelated, they can efficiently be estimated separately. Here, we propose doing so by applying the method of Fan and Yao (1998) to each equation.<sup>1</sup> This method is explained as follows. Suppose the drift function,  $\mu_k(\cdot)$ , is known and let  $R_k = [Y_k/\Delta - \mu_k(X_1, X_2)]^2\Delta$ . Since  $E(R_k|X_1 = x_1, X_2 = x_2) = \sigma_k^2(x_1, x_2)$ , an ideal estimate of the squared diffusion function can be found by regressing  $R_{k,i} = [Y_{k,i}/\Delta - \mu_k(X_{1,i}, X_{2,i})]^2\Delta$  on  $X_{1,i}$  and  $X_{2,i}$ , using LL estimation (see below).

In practice, the drift function is unknown, but we can use the following two-stage procedure. In the first stage, we replace  $\mu_k(X_{1,i}, X_{2,i})$  in  $R_{k,i}$  with an LL estimate,  $\hat{\mu}_k(X_{1,i}, X_{2,i})$ , which is obtained as the value of  $a_0$  solving

$$\min_{a_0, a_1, a_2} \sum_{j=1}^{n-1} [Y_{k,j}/\Delta - a_0 - a_1(X_{1,j} - X_{1,i}) - a_2(X_{2,j} - X_{2,i})]^2 K\left(\frac{X_{1,j} - X_{1,i}}{h_1}\right) K\left(\frac{X_{2,j} - X_{2,i}}{h_2}\right),$$

where  $K(\cdot)$  is a nonnegative bounded kernel function with  $\int K(v)dv = 1$ , and  $h_1$  and  $h_2$  are bandwidths, with  $h_1 \rightarrow 0$ ,  $h_2 \rightarrow 0$ , and  $nh_1h_2 \rightarrow \infty$ , as  $n \rightarrow \infty$ . Popular methods of bandwidth selection include plug-in methods and cross-validation; see Fan and Gijbels (1996).

In the second stage, the residual-based estimator of the squared diffusion function is found by regressing  $\hat{R}_{k,i} = [Y_{k,i}/\Delta - \hat{\mu}_k(X_{1,i}, X_{2,i})]^2\Delta$  on  $X_{1,i}$  and  $X_{2,i}$  using LL estimation. That is, at arbitrary evaluation points  $x_1$  and  $x_2$ , the estimate  $\hat{\sigma}_k^2(x_1, x_2)$  is found as the value of  $b_0$  which solves<sup>2</sup>

$$\min_{b_0, b_1, b_2} \sum_{i=1}^{n-1} [\hat{R}_{k,i} - b_0 - b_1(X_{1,i} - x_1) - b_2(X_{2,i} - x_2)]^2 K\left(\frac{X_{1,i} - x_1}{g_1}\right) K\left(\frac{X_{2,i} - x_2}{g_2}\right),$$

where, for  $k = 1, 2$ ,  $g_k$  is defined analogously to  $h_k$ .

Fan and Yao (1998) refer to this estimator as *adaptive*, as it will estimate the squared diffusion function asymptotically as well as the ideal estimator. We compare the finite-sample performance of these estimators in Section 4.

<sup>1</sup>Fan and Yao (1998) consider only the case where there is a single regressor. Thus, the LL estimator here will have a slower rate of convergence. However, it is still consistent and asymptotically efficient (Ruppert and Wand, 1994; Fan and Gijbels, 1996).

<sup>2</sup>Note that the ideal estimator discussed above is identical to the residual-based estimator, except that  $R_{k,i}$  is used in place of  $\hat{R}_{k,i}$ .

### 3. Specification testing

We now consider specification tests of the general parametric model,

$$\begin{aligned} ds_t &= m_1(s_t, l_t; \theta_1)dt + q_1(s_t, l_t; \phi_1)dW_{1,t} \\ dl_t &= m_2(s_t, l_t; \theta_2)dt + q_2(s_t, l_t; \phi_2)dW_{2,t}, \end{aligned}$$

where, for  $k = 1, 2$ ,  $m_k(\cdot)$ , and  $q_k(\cdot)$  are known functions (which may be linear or non-linear), and  $\theta_k$  and  $\phi_k$  are unknown parameter vectors.

To test the null hypothesis of correct specification, we use the test of Li (1994) and Zheng (1996), henceforth referred to as the  $J_n$  test. Unlike other specification tests used in the literature on continuous-time models, such as the generalized likelihood ratio (GLR) test of Fan et al. (2001), the  $J_n$  test has been shown to be applicable to time-series data (Li, 1999) and insensitive to the choice of bandwidth (Li and Wang, 1998; Kim and Wang, 2006). The  $J_n$  test also benefits from use of the wild bootstrap (Li and Wang, 1998), which is more reliable than the ordinary bootstrap when the model errors are heteroskedastic (Liu, 1988).<sup>3</sup> Application of the  $J_n$  test to the present framework is described below.

#### 3.1. Testing specifications of the drift functions

We first consider testing specifications of the drift functions. That is, for  $k = 1, 2$ , we wish to test

$$H_0^{(k)} : \mu_k(s_t, l_t) = m_k(s_t, l_t; \theta_k) \quad \text{vs.} \quad H_1^{(k)} : \mu_k(s_t, l_t) \neq m_k(s_t, l_t; \theta_k).$$

Note that, in testing specifications of the drift functions, we place no restrictions on the diffusion functions.

Our proposed testing procedure, analogous to that of Kim and Wang (2006), who consider specification tests of the drift function in a single-factor model, is as follows:

**Step 1:** Use least squares to estimate

$$Y_{k,i}/\Delta = m_k(X_{1,i}, X_{2,i}; \theta_k) + u_{k,i}.$$

Call the resulting estimate  $\hat{\theta}_k$ , and let  $\hat{u}_{k,i} = Y_{k,i}/\Delta - m_k(X_{1,i}, X_{2,i}; \hat{\theta}_k)$ .

**Step 2:** Obtain the test statistic,

$$J_n^{(k)} = (n-1)\sqrt{h_1 h_2} I_n^{(k)} / \hat{\Omega}^{(k)},$$

where

$$\begin{aligned} I_n^{(k)} &= \frac{1}{(n-1)} \sum_{i=1}^{n-1} \hat{u}_{k,i} \left[ \frac{1}{(n-2)h_1 h_2} \sum_{j=1, j \neq i}^{n-1} \hat{u}_{k,j} K\left(\frac{X_{1,j} - X_{1,i}}{h_1}\right) K\left(\frac{X_{2,j} - X_{2,i}}{h_2}\right) \right] \\ &= \frac{1}{(n-1)(n-2)h_1 h_2} \sum_{i=1}^{n-1} \sum_{j=1, j \neq i}^{n-1} \hat{u}_{k,i} \hat{u}_{k,j} K\left(\frac{X_{1,j} - X_{1,i}}{h_1}\right) K\left(\frac{X_{2,j} - X_{2,i}}{h_2}\right), \end{aligned}$$

and

$$\hat{\Omega}^{(k)} = \left\{ \frac{2}{(n-1)(n-2)h_1 h_2} \sum_{i=1}^{n-1} \sum_{j=1, j \neq i}^{n-1} \hat{u}_{k,i}^2 \hat{u}_{k,j}^2 \left[ K\left(\frac{X_{1,j} - X_{1,i}}{h_1}\right) K\left(\frac{X_{2,j} - X_{2,i}}{h_2}\right) \right]^2 \right\}^{1/2}.$$

**Step 3:** Construct a bootstrap sample as

$$Y_{k,i}^* = m_k(X_{1,i}, X_{2,i}; \hat{\theta}_k)\Delta + \hat{u}_{k,i}^*,$$

<sup>3</sup>As shown by Li and Wang (1998), even with moderately large samples, the  $J_n$  test statistic (given below) does not come close to attaining its asymptotic distribution (standard normal) under the null.

where  $\{\hat{u}_{k,i}^*\}_{i=1}^{n-1}$  are obtained through a wild bootstrap resampling scheme (see Li and Wang, 1998) from  $\{\hat{u}_{k,i}\}_{i=1}^{n-1}$ .

**Step 4:** Using  $Y_{k,i}^*$  in place of  $Y_{k,i}$ , repeat Steps 1 and 2, calling the test statistic obtained in Step 2  $J_n^{*(k)}$ .

Finally, repeating Steps 3 – 4 some large number of times, the  $p$ -value for the test can be approximated by computing the proportion of times that  $J_n^{*(k)}$  exceeds  $J_n^{(k)}$ .

### 3.2. Testing specifications of the diffusion functions

Analogous to the above, we leave the drift functions unspecified when testing specifications of the diffusion functions. That is, for  $k = 1, 2$ , we test

$$H_0^{(k)} : \sigma_k^2(s_t, l_t) = q_k(s_t, l_t; \phi_k) \quad \text{vs.} \quad H_1^{(k)} : \sigma_k^2(s_t, l_t) \neq q_k(s_t, l_t; \phi_k),$$

while leaving  $\mu_k(\cdot)$  unspecified. This is done as follows:

**Step 1:** For  $i = 1, \dots, (n-1)$ , obtain LL estimates  $\hat{\mu}_k(X_{1,i}, X_{2,i})$ , and let  $\hat{R}_{k,i} = [Y_{k,i}/\Delta - \hat{\mu}_k(X_{1,i}, X_{2,i})]^2 \Delta$ .

**Step 2:** Use least squares to estimate

$$\hat{R}_{k,i} = [q_k(X_{1,i}, X_{2,i}; \phi_k)]^2 + v_{k,i},$$

Call the resulting estimate  $\hat{\phi}_k$ , and let  $\hat{v}_{k,i} = \hat{R}_{k,i} - [q_k(X_{1,i}, X_{2,i}; \hat{\phi}_k)]^2$ .

**Step 3:** Obtain the test statistic,  $J_n^{(k)}$ , as given in Step 2 of Section 3.1, but with  $\hat{v}_{k,i}$  in place of  $\hat{u}_{k,i}$ , and  $g_1$  and  $g_2$  in place of  $h_1$  and  $h_2$ , respectively.

**Step 4:** Construct a bootstrap sample as

$$\hat{R}_{k,i}^* = [q_k(X_{1,i}, X_{2,i}; \hat{\phi}_k)]^2 + \hat{v}_{k,i}^*,$$

where  $\{\hat{v}_{k,i}^*\}_{i=1}^{n-1}$  are obtained through a wild bootstrap resampling scheme from  $\{\hat{v}_{k,i}\}_{i=1}^{n-1}$ .

**Step 5:** Using  $\hat{R}_{k,i}^*$  in place of  $\hat{R}_{k,i}$ , repeat Steps 2 and 3, calling the test statistic obtained in Step 3  $J_n^{*(k)}$ .

Here, the  $p$ -value for the test is approximated by repeating Steps 4 – 5 some large number of times, and computing the proportion of times that  $J_n^{*(k)}$  exceeds  $J_n^{(k)}$ .

## 4. Monte Carlo experiment

In this section, we perform a small Monte Carlo experiment, generating 10,000 samples with  $n = 1,000$  from the parametric model of Schaefer and Schwartz (1984),

$$ds_t = \beta_1(\alpha_1 - s_t)dt + \gamma_1 dW_{1,t} \tag{3}$$

$$dl_t = \beta_2(\alpha_2 - l_t)dt + \gamma_2 \sqrt{l_t} dW_{2,t}. \tag{4}$$

with  $(\beta_1, \alpha_1, \gamma_1) = (0.72, -0.01, 0.007)$  and  $(\beta_2, \alpha_2, \gamma_2) = (0.8587, 0.05, 0.03)$ . Here, for  $k = 1, 2$  we set  $h_k = g_k = c\hat{\sigma}_{X_k} n^{-1/6}$ , where  $c$  is a constant (specified below) and  $\hat{\sigma}_{X_k}$  is the sample standard deviation of  $X_k$ . We employ Gaussian kernel functions.

Table I shows the root mean squared error (RMSE) and mean absolute error (MAE) of the ideal and residual-based estimators of the diffusion functions for  $\Delta = 1/52$  and  $\Delta = 1/250$  (i.e., weekly and daily observations, respectively) and  $c = 1$  (other choices of  $c$  were also used, and the results were qualitatively the same). Overall, the two estimators offer virtually identical performance; the RMSE of the residual-based estimator is consistently just slightly lower than that of the ideal estimator, while the converse is true for the MAE. As expected, daily sampling leads to slightly more accurate estimates than weekly sampling.

Table I: Root mean squared error (RMSE) and mean absolute error (MAE) for the ideal and residual-based estimators of the diffusion functions.

$\Delta$	$k$	Ideal		Residual-based	
		RMSE	MAE	RMSE	MAE
1/52	1	$1.157 \times 10^{-5}$	$7.686 \times 10^{-6}$	$1.091 \times 10^{-5}$	$7.972 \times 10^{-6}$
	2	$1.087 \times 10^{-5}$	$7.106 \times 10^{-6}$	$1.024 \times 10^{-5}$	$7.374 \times 10^{-6}$
1/250	1	$1.019 \times 10^{-5}$	$7.144 \times 10^{-6}$	$0.975 \times 10^{-5}$	$7.322 \times 10^{-6}$
	2	$0.943 \times 10^{-5}$	$6.570 \times 10^{-6}$	$0.902 \times 10^{-5}$	$6.744 \times 10^{-6}$

To assess the size of the  $J_n$  test, we test the null hypotheses that the drift and diffusion functions are of the form specified in (3) – (4). In order to reduce the computational cost, we skip Step 1 in testing the diffusion functions, and assume that the drift functions are known (i.e., we use the ideal squared residuals,  $R_{k,i}$ ). Here, we focus on the weekly case (i.e.,  $\Delta = 1/52$ ), but, to better understand the consequences of bandwidth choice, we use  $c = 0.5$ ,  $c = 1.0$ , and  $c = 1.5$ . Rejection rates at the 5% level, based on 199 bootstrap replications, are shown in Table II.

Table II: Rejection rates for specification tests of the drift and diffusion functions.

$c$	$k$	Drift	Diffusion
0.5	1	0.0487	0.0514
	2	0.0514	0.0512
1.0	1	0.0527	0.0499
	2	0.0511	0.0515
1.5	1	0.0529	0.0499
	2	0.0524	0.0496

From Table II, it is clear that the actual rejection rates, for both the drift and diffusion functions, are quite close to the nominal level for all values of  $c$ . This confirms the results of Li and Wang (1998) and Fan and Wang (2006), who find that the  $J_n$  test is not sensitive to the choice of bandwidth.

## 5. Empirical example

We close by applying our proposed methodology to some U.S. Treasury security data, which consists of  $n = 1,444$  weekly observations, covering the period January 2, 1981 to August 28, 2009. The 3-month secondary market rate is used as the short rate, while the 10-year constant maturity rate is used as the consol rate. Bandwidths are chosen using least-squares cross-validation, and, as above, Gaussian kernel functions are used.

Nonparametric estimates of the two-factor model in (1) – (2) are plotted in Figure 1. For the parametric specification in (3) – (4), the  $J_n$  test statistics are 1.3205, 2.9309, 13.2746, and 10.4073, for the spread drift, consol drift, spread diffusion, and consol diffusion, respectively. Using 999 bootstrap replications, we estimate  $p$ -values for these statistics of 0.0561, 0.0050, 0, and 0, respectively. Hence, we can reject all of the restrictions at the 1% level, except that on the spread drift function, which can only be rejected at the 10% level. Thus, this finding can be seen to offer some limited support for the method of Knight et al. (2006), which places a linear restriction on the spread drift function. That being said, our proposed estimation method does offer some increased flexibility, as such a restriction may be more clearly rejected in other datasets.

## References

- Chapman, D.A. and N.D. Pearson (2000) "Is the short rate drift actually non-linear?" *Journal of Finance* **55**, 355-388.
- Fan, J. and I. Gijbels (1996) *Local Polynomial Modelling and Its Applications*, Chapman & Hall: London.
- Fan, J. and Q. Yao (1998) "Efficient estimation of conditional variance functions in stochastic regression" *Biometrika* **85**, 645-660.

- Fan, J. and C. Zhang (2003) "A reexamination of diffusion estimators with applications to financial model validation" *Journal of the American Statistical Association* **98**, 118-134.
- Fan, J., C. Zhang, and J. Zhang (2001) "Generalized likelihood ratio statistics and Wilks phenomenon" *Annals of Statistics* **29**, 153-193.
- Kim, M.S. and S. Wang (2006) "Sizes of two bootstrap-based nonparametric specification tests for the drift function in continuous time models" *Computational Statistics & Data Analysis* **50**, 1793-1806.
- Knight, J.L., F. Li, and M. Yuan (2006) "A semiparametric two-factor term structure model" *Journal of Financial Econometrics* **4**, 204-237.
- Li, Q. (1994) "A consistent test for linearity in partially linear regression models" University of Guelph discussion paper number 1994-7.
- Li, Q. (1999) "Consistent model specification tests for time series econometric models" *Journal of Econometrics* **92**, 101-147.
- Li, Q. and S. Wang (1998) "A simple consistent bootstrap test for a parametric regression function" *Journal of Econometrics* **87**, 145-165.
- Liu, R.Y. (1988) "Bootstrap procedures under some non-I.I.D. models" *Annals of Statistics* **16**, 1696-1708.
- Ruppert, D. and M.P. Wand (1994) "Multivariate weighted least squares regression" *Annals of Statistics* **22**, 1346-1370.
- Schaefer, S.M. and E.S. Schwartz (1984) "A two-factor model of the term structure: An approximate analytical solution" *Journal of Financial and Quantitative Analysis* **19**, 413-424.
- Zheng, J.X. (1996) "A consistent test of functional form via nonparametric estimation technique" *Journal of Econometrics* **75**, 263-289.

Figure 1: Nonparametric estimates of the drift and diffusion functions

