Diffusion models for exchange rates in a target zone.

Kristian Stegenborg Larsen
Nykredit, Kalvebod Brygge 1–3
DK-1780 Copenhagen V
Denmark

Michael Sørensen
Department of Applied Mathematics and Statistics
University of Copenhagen
Universitetsparken 5, DK-2100 København Ø
Denmark

Abstract

We present two relatively simple and analytically tractable diffusion models for an exchange rate in a target zone. One model generalize a model proposed by De Jong, Drost & Werker (2001) to allow asymmetry between the currencies and thus obtain a better fit to data, in which asymmetries are often an important feature. Optimal estimation of the model parameter using eigenfunctions of the generator is investigated in detail and shown to give well-behaved estimators that are easy to calculate. The model is demonstrated to fit data on exchange rates in the European Monetary System well. Also an alternative diffusion model is presented, which has similar properties in the centre of the target zone, but with a more realistic dynamics near the boundaries of the target zone. Estimators based on eigenfunctions work well in this case too. For both models no-arbitrage pricing of derivative assets is discussed. Finally, problems concerning adjustments of the central parity are discussed.

Key words: currency options, eigenfunctions, estimating functions, exchange rate target zones, Jacobi diffusion, option pricing, realignments.

1 Introduction

We propose and investigate two relatively simple and analytically tractable target zone models, i.e. models for an exchange rate which the central banks involved have promised to keep within fixed bounds. A much studied example of target zones is the European Monetary System (EMS) that was in place in some form or another from 1979 until the introduction of the euro on January 1, 1999. For a number of European currencies a target zone for the exchange rate with the euro has been arranged, formally or informally. Our first model generalize the target zone model proposed by De Jong, Drost & Werker (2001) by allowing asymmetry between the currencies, because asymmetry is an important feature of data on exchange rates in a target zone. We also propose an alternative model, which has a behaviour quite similar to the first model in the central part of the target zone, but with a very different, and probably more realistic, dynamics near the boundaries of the target zone.

Since the seminal paper by Krugman (1991), a large amount of research on target zones has been reported in the econometric literature; for a short review, see Christensen, Lando & Miltersen (1998) or De Jong, Drost & Werker (2001). In both of our models, the intervention policy of the central banks and the behaviour of speculators is modelled implicitly by specifying the drift and diffusion coefficients. Models that more explicitly describe the intervention policy and the speculator behaviour may be seen as preferable because their economic interpretation is more direct. However, results based on such models may depend crucially on details of the specification, and the models may be difficult to handle in practice. Advantages of the models presented here are that they are flexible enough to fit data on exchange rates in target zones well, but on the other hand are simple and tractable enough that estimation of the parameters can be done easily and as efficiently as desired. Also option pricing is easy.

We present the first model, the Jacobi diffusion, in the Section 2. For this diffusion the eigenfunctions of the infinitesimal generator are explicitly know: They are essentially the Jacobi polynomials. The fact that the eigenfunctions are known renders it natural and simple to estimate the parameters of the model by means of an optimal estimating function based on the eigenfunctions as proposed by Kessler & Sørensen (1999). These estimators are easy to calculate as they are solutions to an explicit system of equations. The rest of the section investigates the large sample performance of this relatively new statistical technique in the case of the Jacobi diffusions. This is of independent statistical interest apart from the financial context and indicates that estimators based on just two eigenfunctions can be surprisingly efficient. The properties of the estimators in finite samples are studied in a simulation study in Section 3. For some parameters the asymptotic results can be used for relatively small samples of 300 observations, but at least for one parameter, the speed of mean-reversion, considerably more observations are needed before the asymptotic theory can be used. When this happens simulation is a useful alternative. In Section 4 weekly observations of the exchange rates of the German mark to the Danish krone, the French franc, the Belgian franc, and the Dutch guilder are analyzed by means of the Jacobi diffusion model. The data are from a period in the time of the European Monetary System where no adjustments of the central parity took place. The fit of the model to the data is investigated in a number of ways and found to be good. In order to obtain a better fit for data points close to the boundaries of the target zone, an alternative model is proposed in Section 5. In the centre of the target zone, the two models have a similar behaviour. The new model is related to the Jacobi diffusion by a simple transformation, and is hence as tractable as the Jacobi diffusion. In particular, estimation can be done easily by means of an optimal estimating function based on the explicit eigenfunctions. In Section 6 no-arbitrage pricing of options and other contingent claims on the exchange rate is considered. It is demonstrated that care must the exercised when modelling the relation between the domestic and the foreign interest rates in order to avoid a model with arbitrage opportunities. In the previous sections the risk of adjustments to the central parity was ignored. This difficult problem is beyond the scope of the present paper, but is discussed at some length in Section 7, where also the statistical problems in connection with a two-factor model are briefly considered. Finally, the Appendix explains how to explicitly calculate the optimal estimating functions needed to estimate the parameters of the two models.

2 The Jacobi diffusion

Let S_t denote the exchange rate at time t, and suppose that it has been agreed that the exchange rate may not be more than $a \cdot 100$ per cent over the central parity μ for any of the two currencies, i.e. $\mu/(1+a) < S_t < \mu(1+a)$. Then $X_t = \log(S_t)$ must satisfy that $m - z < X_t < m + z$, where $m = \log(\mu)$ and $z = \log(1+a)$.

As a model for the logarithm of the exchange rate X_t, we propose the diffusion model defined by

$$dX_t = -\beta[X_t - (m + \gamma z)]dt + \sigma\sqrt{z^2 - (X_t - m)^2}dW_t, \tag{2.1}$$

where $\beta > 0$, $\gamma \in (-1, 1)$, and W is a standard Wiener process. Clearly, this process reverts to the mean $m + \gamma z$ with a speed that is proportional to the deviation from this level. The parameter γ is an asymmetry parameter that express whether one currency is stronger than the other. When the process gets near the boundaries $m - z$ or $m + z$, the diffusion coefficient becomes small and the drift (which models the intervention of the central banks) drives the process away from these boundaries. The drift at the upper boundary $m + z$ is $-\beta(1 - \gamma)z$, while the drift at the lower boundary $m - z$ is $\beta(1 + \gamma)z$. The process (2.1) is a particular case of the general class of mean-reverting diffusion processes with a given marginal distribution studied in Bibby, Skovgaard & Sørensen (2003). The symmetric case $\gamma = 0$ is the target zone model proposed by De Jong, Drost & Werker (2001). Jacobi diffusions have also been studied by Gourieroux & Jasiak (2003), who introduced a multivariate version and presented several applications.

Easy calculations involving the speed and scale measure shows that the diffusion is an ergodic diffusion on the interval $(m - z, m + z)$ if and only if $\kappa_1 = \beta(1 - \gamma)\sigma^{-2} \geq 1$ and $\kappa_2 = \beta(1 + \gamma)\sigma^{-2} \geq 1$, or otherwise expressed $\beta \geq \sigma^2$ and $-1 + \sigma^2/\beta \leq \gamma \leq 1 - \sigma^2/\beta$. The invariant distribution is a shifted and rescaled Beta-distribution on the interval $(m - z, m + z)$ with parameters κ_1 and κ_2, i.e. the distribution with probability density

$$\pi(x) = (z + m - x)^{(\kappa_1 - 1)}(z - m + x)^{(\kappa_2 - 1)}(2z)^{(1 - \kappa_1 - \kappa_2)}/B(\kappa_1, \kappa_2), \quad x \in (m - z, m + z), \tag{2.2}$$

where B denotes the beta-function. Thus if the distribution of X_0 is (2.2), then the solution to (2.1) is stationary. If $\beta(1 - \gamma)\sigma^{-2} < 1$, then X can reach the boundary $m + z$ at a finite time point, and if $\beta(1 + \gamma)\sigma^{-2} < 1$, the boundary $m + z$ can be reached in finite time. The economic interpretation is that the intervention of the central banks must be sufficiently forceful compared to both the marked volatility and the asymmetry between the currencies to keep the process stationary.

The eigenfunctions of the generator of the diffusion (2.1) are explicitly known and are given by

$$\varphi_n(x; \beta, \gamma, \sigma, m, z) = P_n^{(\beta(1-\gamma)\sigma^{-2}-1,\, \beta(1+\gamma)\sigma^{-2}-1)}((x - m)/z), \tag{2.3}$$

where $P_n^{(a,b)}(x)$ denotes the Jacobi polynomial of order n given by

$$P_n^{(a,b)}(x) = \sum_{i=0}^{n} 2^{-i} \binom{n+a}{n-i} \binom{a+b+n+i}{i} (x-1)^i, \tag{2.4}$$

see Abramowitz & Stegun (1964). The eigenvalue corresponding to φ_n is

$$\lambda_n(\beta,\gamma,\sigma) = n\left(\beta + \tfrac{1}{2}\sigma^2(n-1)\right). \tag{2.5}$$

For estimation of the parameters of the model, it is therefore natural to use one of the optimal estimating functions based on the eigenfunctions of the generator proposed by Kessler & Sørensen (1999). This is particularly appealing in this case because for diffusions on a finite interval it was proved in Kessler & Sørensen (1999) that by including sufficiently many eigenfunctions, an estimator can be obtained that is as close to being efficient as one wishes.

Assume that the process X has been observed at discrete time points, $X_{t_0}, X_{t_1}, \ldots, X_{t_n}$ ($t_0 = 0$). If the target zone has been officially declared, the quantities m and z are known. This is the case for the European Monetary System. In some cases target zones are not declared, or the actual target zone is different from the official target zone. In such cases m and z are parameters to be estimated. In the following θ denotes either (β,γ,σ) or $(\beta,\gamma,\sigma,m,z)$. An estimating function based on the first N eigenfunctions is given by

$$G_n(\theta) = \sum_{i=1}^{n} A(\Delta_i, X_{t_{i-1}}; \theta) h(\Delta_i, X_{t_{i-1}}, X_{t_i}; \theta), \tag{2.6}$$

where $\Delta_i = t_i - t_{i-1}$, $A(\Delta, x; \theta)$ is a $p \times N$-matrix, p is the dimension of θ, and $h = (h_1, \ldots, h_N)^T$ with

$$h_j(\Delta, x, y; \theta) = \varphi_j(y;\theta) - e^{-\lambda_j(\theta)\Delta}\varphi_j(x;\theta), \tag{2.7}$$

$j = 1, \ldots, N$. Conditions in Kessler & Sørensen (1999) ensuring that (2.6) is a martingale when θ is the true parameter value are satisfied, because $x \mapsto \sqrt{z^2 - (x-m)^2}$ and $\partial_x \varphi_j(x)$ are bounded on the state space. The optimal choice in the sense of Godambe & Heyde (1987) (see also Heyde (1997)) of the weight matrix $A(\Delta, x; \theta)$ can be found explicitly, see the Appendix. Thus the estimator of θ is given as the solution to an explicit system of p equation. Kessler & Sørensen (1999) showed that for N going to infinity the optimal estimating function of the type (2.6) will converge to the score function.

We shall mainly be concerned with the case where $\theta = (\beta,\gamma,\sigma)$. In this case the conditions in Kessler & Sørensen (1999) ensuring (asymptotic) existence, consistency and asymptotic normality of the estimator $\hat{\theta}_n$ obtained by solving the estimating equation $G_n(\theta) = 0$ can be shown to hold provided that $N > 1$. The expression for the asymptotic covariance matrix of $\hat{\theta}_n$ when the optimal estimating function is used can be found in the Appendix.

Let us discuss the asymptotic variance for different choices of the eigenfunctions used to define (2.6). It was decided to use in this discussion the parametrization $(\alpha_1, \alpha_2, \sigma^2)$, where the parameters $\alpha_1 = \beta(1-\gamma)$ and $\alpha_2 = \beta(1+\gamma)$ determine the drift away from the upper and the lower boundary, respectively. One could alternatively have considered the parameters κ_1 and κ_2 that determine the invariant distribution, but in a study of asymptotic variance it is advisable to choose a parametrization that separates parameters in the drift and parameters in the diffusion coefficient as the asymptotic variances of these two types of parameters can be quite different when Δ is small. We shall see that this actually is the case. To simplify the calculations we will consider the asymptotic information (inverse of asymptotic variance) for each of the parameters when the other parameters are fixed. These quantities are the diagonal elements of the matrix $I(\alpha_1, \alpha_2, \sigma^2)$ in the Appendix. The information quantities in the following tables were calculated by means of Maple. We consider two parameter values $(\alpha_1, \alpha_2, \sigma^2) = (0.02, 0.02, 0.01)$ and $(\alpha_1, \alpha_2, \sigma^2) = (0.06, 0.04, 0.02)$, which are of the same order of magnitude as the estimates we obtain in Section 4 from EMS data. In both cases we take $m = 0$ and $z = 1$ and the sampling points to be equidistant with $t_i - t_{i-1} = \Delta = 1$.

First we consider the case where the estimating function is based on only one eigenfunction. This need not be the first eigenfunction. The information quantities are calculated for the eigenfunctions based on φ_j, $j = 1, \ldots, 8$. The results are given in the Tables 2.1 and 2.2. When $\alpha_1 = \alpha_2$, it is not difficult to see that the information for these two parameters must be the same. That the information about σ^2 is zero for φ_1, is due to the fact that this parameter cannot be estimated based on a linear estimating function.

Eigenfunction no.	1	2	3	4	5	6	7	8
Inf. for $\hat{\alpha}_1$ and $\hat{\alpha}_2$	47.7	44.8	41.6	38.7	36.1	33.6	31.2	28.6
Inf. for $\hat{\sigma}^2$	0.0	758.7	1103.4	1346.8	1538.6	1692.1	1810.3	1892.6

Table 2.1: Information for the estimating functions based on one eigenfunction and with $(\alpha_1, \alpha_2, \sigma^2) = (0.02, 0.02, 0.01)$.

Eigenfunction no.	1	2	3	4	5	6	7	8
Inf. for $\hat{\alpha}_1$	12.2	10.4	9.0	8.0	7.0	6.0	5.0	4.0
Inf. for $\hat{\alpha}_2$	34.1	31.7	28.7	25.8	22.9	20.0	16.9	13.9
Inf. for $\hat{\sigma}^2$	0.0	287.0	390.0	452.0	489.7	505.3	498.5	469.9

Table 2.2: Information for the estimating functions based on one eigenfunction and with $(\alpha_1, \alpha_2, \sigma^2) = (0.06, 0.04, 0.02)$.

We see that for α_1 and α_2 the information is largest for the first eigenfunction in both cases. For σ^2 the maximum information is obtained with the 8th and the 6th eigenfunction. As one would expect, it is necessary to combine eigenfunctions to get a good estimating function for all parameters. In the Tables 2.3 and 2.4 are given the information quantities for the estimating functions based on combinations of two of the eigenfunctions φ_j, $j = 1, \ldots, 5$.

	Inf. for $\hat{\alpha}_1$ and $\hat{\alpha}_2$				Inf. for $\hat{\sigma}^2$			
	2	3	4	5	2	3	4	5
1	49.2	49.2	49.3	49.2	5016	4467	4129	3866
2		49.2	46.2	49.2		5037	4122	4319
3			49.1	49.0			5034	4456
4				48.8				5017

Table 2.3: Information for the estimating functions based on two eigenfunctions and with $(\alpha_1, \alpha_2, \sigma^2) = (0.02, 0.02, 0.01)$. The indexes of the eigenfunctions are indicated.

	Inf. for $\hat{\alpha}_1$				Inf. for $\hat{\alpha}_2$				Inf. for $\hat{\sigma}^2$			
	2	3	4	5	2	3	4	5	2	3	4	5
1	12.5	12.4	12.4	12.4	36.1	36.2	36.2	36.2	1264	1061	945	852
2		12.4	11.9	12.2		36.1	35.2	36.0		1264	1046	1005
3			12.1	11.9			35.6	35.3			1254	1039
4				11.7				34.5				1233

Table 2.4: Information for the estimating functions based on two eigenfunctions and with $(\alpha_1, \alpha_2, \sigma^2) = (0.06, 0.04, 0.02)$. The indexes of the eigenfunctions are indicated.

We see only a minor increase in the information about α_1 and α_2, but a dramatic increase for σ^2. Based on Tables 2.3 and 2.4 there is no reason to use anything but the simplest estimating function based on two eigenfunctions, i.e. that based on the first two eigenfunctions. We have also calculated the information quantities for estimating functions based on a larger number of eigenfunctions, but it does not seem possible to increase the information quantities in this way by more than 1 - 3 per cent. As mentioned the estimators will tend to be efficient as the number of eigenfunctions goes to infinity, so this is a strong indication that estimators based on the two first eigenfunctions are close to being efficient, at least for the two parameter values considered in this study.

3 A simulation study

In this section we report a simulation study to investigate the finite sample properties of the estimators obtained from the optimal estimating function based on the first two eigenfunctions. The process (2.1) was simulated by means of the strong Taylor scheme of order 1.5, see Kloeden & Platen (1999). We consider equidistant observation times separated by $\Delta = 1$. The discretization used for the simulation was much finer: The time between the simulation time points was 0.0001. Before the first observation was made, the process was simulated for 50 time units.

Let us first mention some general experiences obtained in the simulation study. It turned out that the numerical problems in finding the solution to the estimating equations were least when the estimating function was multiplied by σ^2, and when the parametrization $(\log(\beta), \gamma, \log(\sigma^2))$ was used in the numerical search procedure. In the previous section we saw that the information about the parameters α_1 and α_2 (and hence β and γ) is quite high when the optimal estimating function based only on the first eigenfunction is used. Moreover, it has turned out that estimates of β and γ obtained from this estimating function depend only very little on the value of σ^2. Therefore the following procedure was used and turned out to work very well. First preliminary estimates of β and γ obtained from the optimal estimating function based on the first eigenfunction with σ^2 fixed at some reasonable value. Then a preliminary estimate of σ^2 was obtained from the optimal estimating function based on the first two eigenfunctions with β and γ fixed at their preliminary estimates using bisection. Finally, the estimates of all three parameters obtained from the optimal estimating function based on the first two eigenfunctions was found by means of Broyden's method with the preliminary estimates used as starting values. For details about Broyden's method and its implementation, see Press et al. (1992).

First the process (2.1) was simulated with the parameter values $\beta = 0.05$, $\gamma = -0.2$ and $\sigma^2 = 0.02$, for which the process is ergodic. For 250 independent time series of 300 simulated observations, 100 series of 600 observations, and 50 series of 1200 observations the three parameters were estimated, and for each sample size the mean and variance of the simulated estimators were calculated. The results are given in Table 3.1.

# obs.	# sim.	mean $\hat{\beta}_n$	SE $\hat{\beta}_n$	mean $\hat{\gamma}_n$	SE $\hat{\gamma}_n$	mean $\hat{\sigma}_n^2$	SE $\hat{\sigma}_n^2$
300	250	.0645	.0229	-.195	.148	.02015	.00184
600	100	.0547	.0117	-.195	.099	.02005	.00122
1200	50	.0516	.0077	-.193	.071	.02001	.00087

Table 3.1: Mean and standard error of the estimators of β, γ and σ^2. The true parameter values are β=0.05, γ=-0.2, and σ^2=0.02.

The estimators of γ and σ^2 are found to be almost unbiased, while β is somewhat overestimated when there are only 300 observations, but is well estimated with 1200 observations. It is well known that the speed of reversion is relatively difficult to estimate without bias. For γ and σ^2 the standard error times the square-root of the sample size is almost independent of the sample size, and for σ^2 this quantity is close to the value obtained from the information calculated in Section 2. This indicates that the asymptotics work already with a sample size of 300. For β the asymptotics clearly does not work with only 300 observations, but appears to work well with 1200 observations. To investigate this further, normal quantile plots have been drawn for the simulated estimators, see Figure 3.1. These plots are in accordance with the conclusion just made: There seems to be serious problems with the normality only for $\hat{\beta}_n$ with sample size $n = 300$ and perhaps $n = 600$. Finally, the correlation between the estimators $\hat{\beta}$ and $\hat{\gamma}$ was estimated and found to be small (between -0.1 and -0.17). Also the correlation between $\hat{\alpha}_1$ and $\hat{\alpha}_2$ was estimated and found to be much larger (about 0.6 in all cases). This indicates that the parametrization $(\beta, \gamma, \sigma^2)$ is to be preferred over $(\alpha_1, \alpha_2, \sigma^2)$ for statistical inference.

Next we consider simulations of 48 time series of 1200 observations, 96 series of 600 observations and 242 series of 300 observations. The parameter values are now $\beta = 0.02$, $\gamma = 0$ and $\sigma^2 = 0.01$, so again the process is ergodic. The results are given in Table 3.2. The tendencies are as in Table 3.1, but the bias of $\hat{\beta}_n$ is somewhat larger (note the smaller standard errors).

Finally, we consider simulations with the parameter values $(\beta, \gamma, \sigma^2) = (0.02, 0, 0.02)$ and $(\beta, \gamma, \sigma^2) = (0.02, 0, 0.05)$, see Tables 3.3 and 3.4. For the latter set of parameter values the process is not ergodic, and both boundaries have a positive probability of being hit at a finite time, and the asymptotics cannot be expected to work. However, in both cases the pattern is the same as in the Tables 3.1 and 3.2.

# obs.	# sim.	mean $\hat{\beta}_n$	SE $\hat{\beta}_n$	mean $\hat{\gamma}_n$	SE $\hat{\gamma}_n$	mean $\hat{\sigma}_n^2$	SE $\hat{\sigma}_n^2$
300	242	.0351	.0165	.017	.239	.01012	.00081
600	96	.0253	.0079	.010	.156	.01008	.00057
1200	48	.0224	.0042	.013	.108	.01006	.00039

Table 3.2: Mean and standard error of the estimators of β, γ and σ^2. The true parameter values are β=0.02, γ=0, and σ^2=0.01.

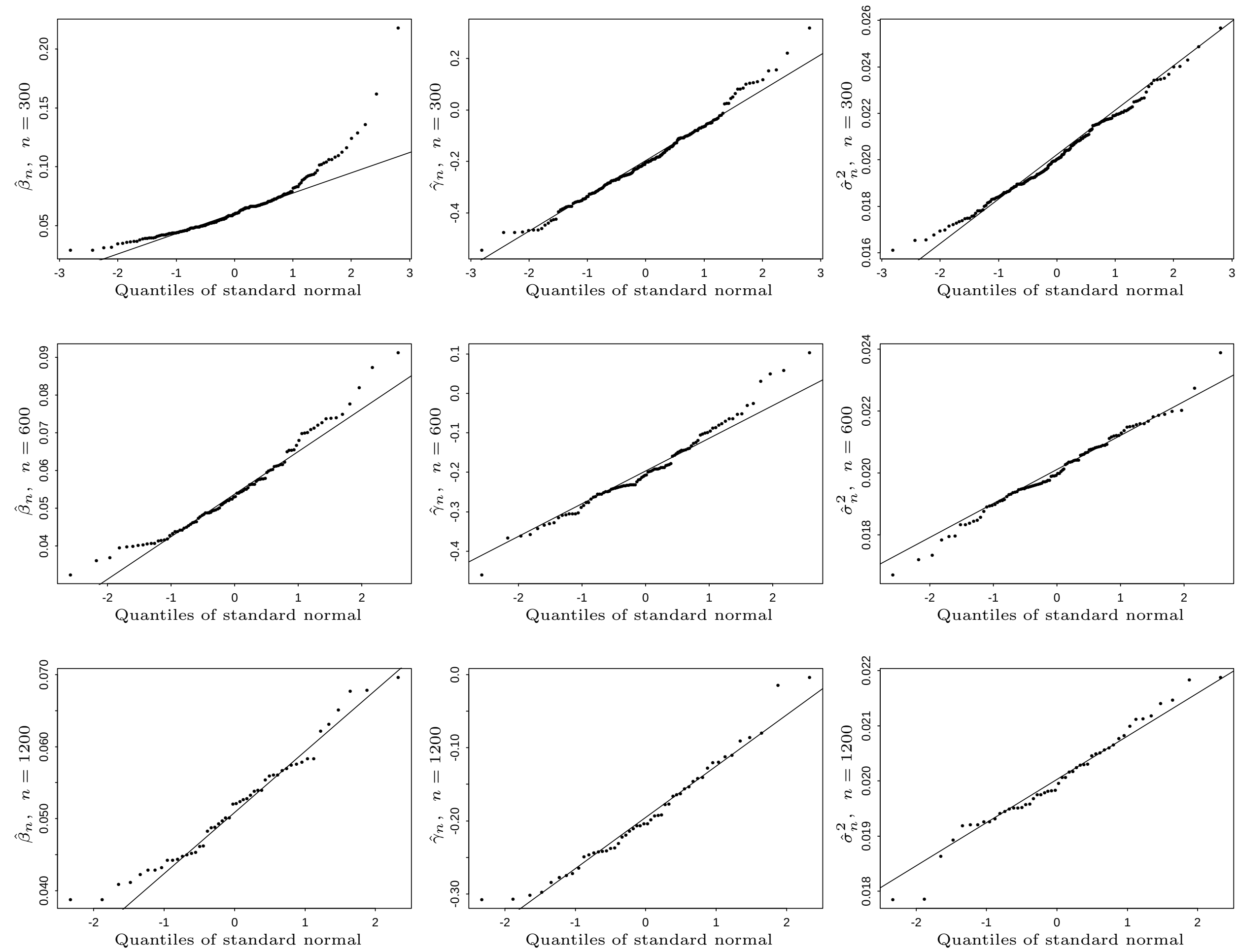

Figure 3.1: Normal quantile plots for the simulated estimators $\hat{\beta}_n$, $\hat{\gamma}_n$, and $\hat{\sigma}^2_n$ with the sample size n equal to 300, 600, and 1200.

# obs.	# sim.	mean $\hat{\beta}_n$	SE $\hat{\beta}_n$	mean $\hat{\gamma}_n$	SE $\hat{\gamma}_n$	mean $\hat{\sigma}^2_n$	SE $\hat{\sigma}^2_n$
300	200	.0342	.0155	.006	.312	.0199	.0018
600	100	.0266	.0086	.003	.234	.0198	.0012
1200	50	.0229	.0061	-.010	.155	.0198	.0009

Table 3.3: Mean and standard error of the estimators of β, γ and σ^2. The true parameter values are β=0.02, γ=0, and σ^2=0.02.

# obs.	# sim.	mean $\hat{\beta}_n$	SE $\hat{\beta}_n$	mean $\hat{\gamma}_n$	SE $\hat{\gamma}_n$	mean $\hat{\sigma}^2_n$	SE $\hat{\sigma}^2_n$
300	200	.0319	.0156	-.015	.336	.0499	.0047
600	100	.0247	.0085	-.014	.188	.0498	.0031
1200	50	.0224	.0024	-.024	.096	.0498	.0019

Table 3.4: Mean and standard error of the estimators of β, γ and σ^2. The true parameter values are β=0.02, γ=0, and σ^2=0.05.

4 The EMS data

In this section we estimate the parameters of the model based on time series of exchange rates and investigate how well the model fits these data. The data are weekly observations of the exchange rates of the German mark to the Danish krone, the French franc, the Belgian franc, and the Dutch guilder in 298 weeks from January 1987 till August 1992. In this period the central parity μ was not changed, and the largest allowed deviation was 2.25 per cent. Figures 4.1 – 4.4 show the observed values of $X_t = 100[\log(S_t) - \log(\mu)]$, where S_t denotes the exchange rate at time t. We will model the process X by (2.1) with $\mu = 0$ and $z = 2.25$.

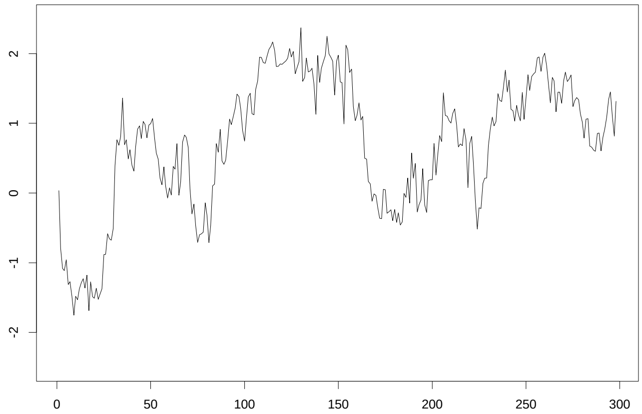

Figure 4.1: The Danish krone.

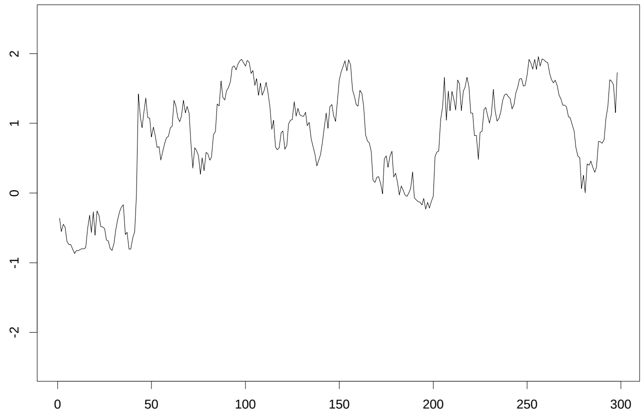

Figure 4.2: The French franc.

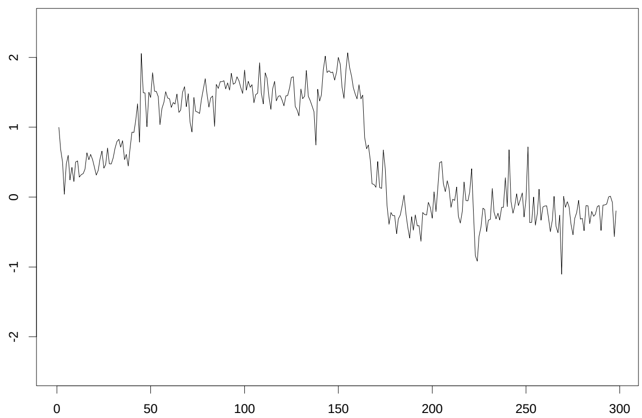

Figure 4.3: The Belgian franc.

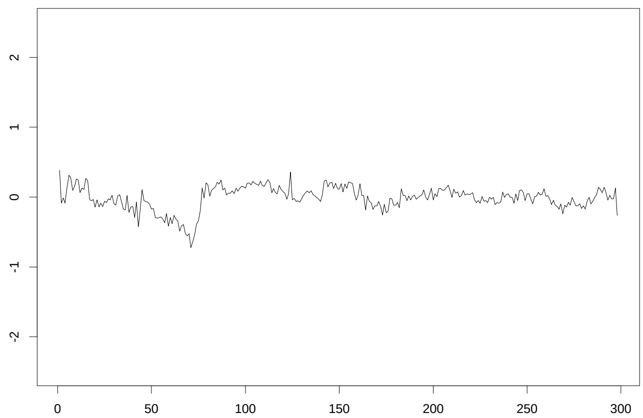

Figure 4.4: The Dutch guilder.

Estimates of the parameters β, γ and σ^2 are found by the procedure described in Section 3 and are given in Table 4.1. The estimates for the Danish, French and Belgian currencies are similar, in particular for the Danish krone and the French franc, whereas the estimates for the Dutch guilder are quite different. For the Dutch guilder the speed of reversion is larger, the random variation (the diffusion) is smaller, and the data are symmetric. For all the estimated parameter vectors the process (2.1) is ergodic.

In order to get an idea about how well the parameters have been determined, a small

	$\hat{\beta}_n$	$\hat{\gamma}_n$	$\hat{\sigma}_n^2$
Danish krone	.0467	.319	.0185
French franc	.0474	.399	.0117
Belgian franc	.0755	.262	.0207
Dutch guilder	.2131	-.011	.0023

Table 4.1: Estimates for the 4 currencies.

simulation study was performed. For each of the estimated parameter vectors, 50 independent time series of 300 observations each were simulated and the parameters estimated. In this way standard errors were obtained for the estimates, and confidence intervals were calculated using these standard errors and assuming a normal distribution, see Table 4.2. At least in the case of the parameter β, the normal distribution is a rather dubious assumption in view of the results in Section 3, but the intervals in Table 4.2 do give an impression of the precision of the estimates. It can safely be concluded that the speed of mean reversion parameter β is significantly higher for the Dutch guilder than for the Danish and French currencies. It is also quite clear that the asymmetry parameter γ is positive for the Danish, French and Belgian currencies, while it can safely be assumed to be zero for the Dutch guilder. On the basis of the results in Section 3 it must be expected that the speed of mean reversion is over-estimated. The small simulation study confirmed this.

	β	γ	σ^2
Danish krone	(.010;.083)	(.070;.568)	(.0154;.0216)
French franc	(.011;.084)	(.176;.622)	(.0095;.0139)
Belgian franc	(.020;.131)	(.089;.434)	(.0168;.0246)
Dutch guilder	(.121;.305)	(-.036;.014)	(.0019;.0027)

Table 4.2: 95% confidence intervals.

Let us next investigate how well the Jacobi diffusion (2.1) fits the data. In the first plot of the Figures 4.5 – 4.8, we compare the marginal beta distributions with the estimated parameter values to histograms of the observations X_{t_i} for the four currencies. The second plot of the figures are quantile plots for the estimated beta distributions. The marginal distributions appear to be well fitted by the estimated beta distributions for the Danish krone, the French franc and the Dutch guilder, perhaps with small problems at the lower tail of the distribution for the Danish krone and the French franc and with a kink in the lower half of the quantile plot for guilder. A simulation study showed that the small deviations seen for the Danish krone and the French franc can be easily explained by sample randomness. For the Belgian franc the marginal distribution seems to be bimodal and is thus not well modelled by a beta distribution. The data in Figure 4.3 indicate that a regime shift took place after week number 150. A likely explanation is that a non-declared adjustment of the central parity took place at that time. It is possible that the model would fit the two regimes separately, but this has not been investigated. Note that in all cases, apart from the Dutch guilder, it is an important condition for the good fit that the model allows a non-symmetric beta distribution.

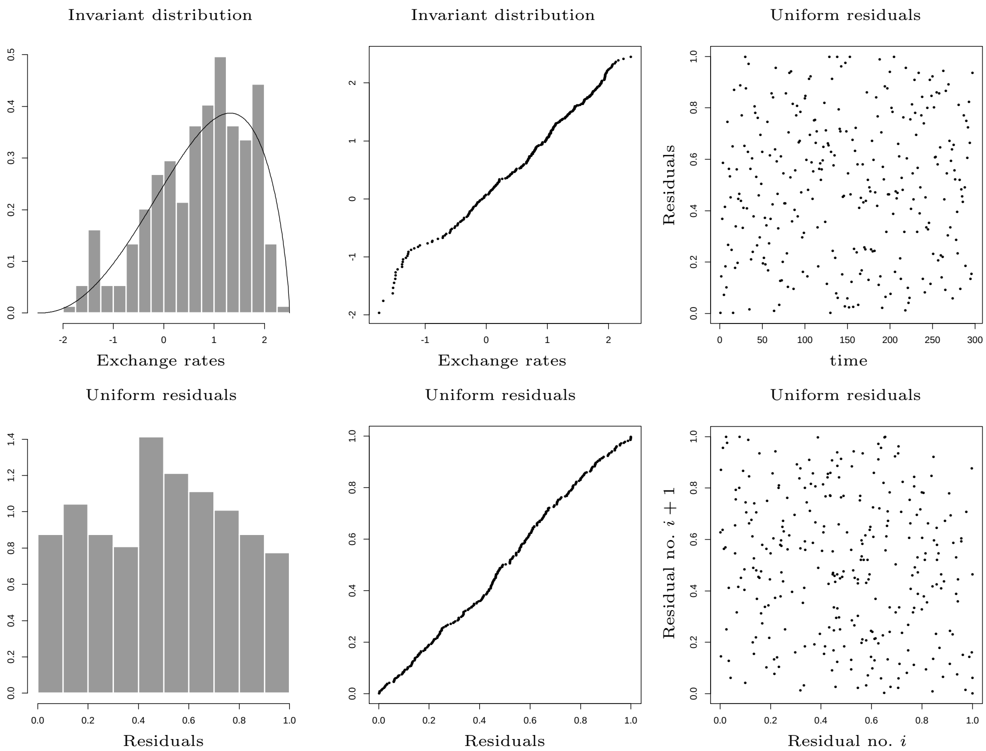

Figure 4.5: The Danish krone.

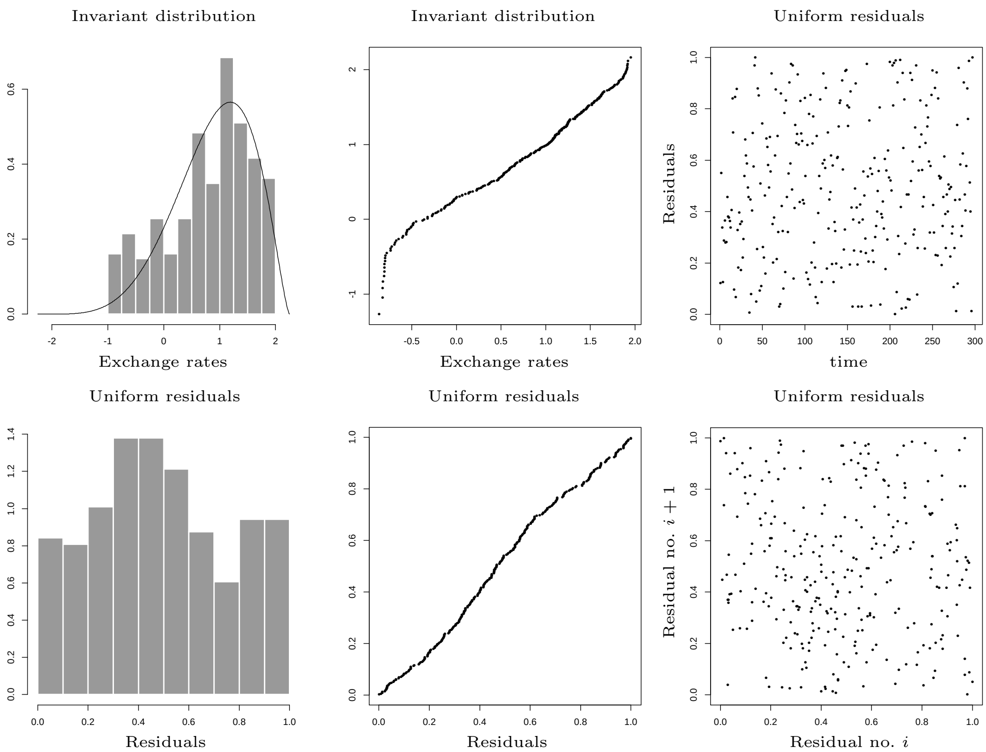

Figure 4.6: The French franc.

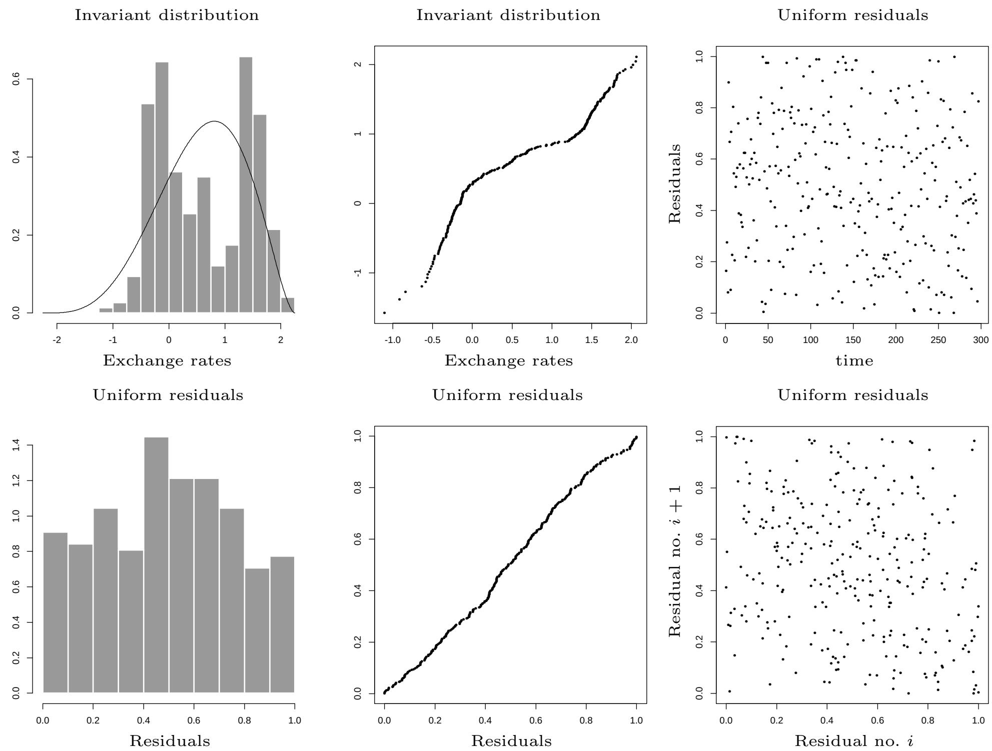

Figure 4.7: The Belgian franc.

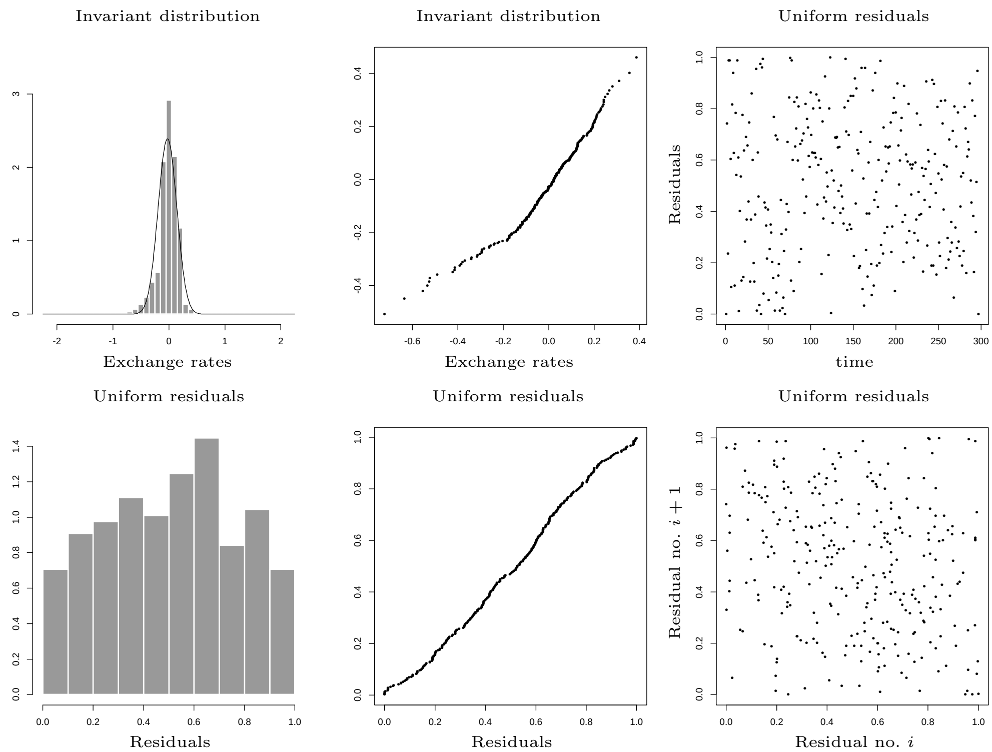

Figure 4.8: The Dutch guilder.

To investigate the fit of the transition distribution, which is not explicitly known, we use the simulated uniform residuals proposed by Pedersen (1994). These are given by $U_i = F(X_{t_i}|X_{t_{i-1}};\beta,\gamma,\sigma^2)$, $i = 1,\dots,297$, where $F(y|x;\beta,\gamma,\sigma^2)$ denotes the conditional distribution function of X_{t_1} given that $X_0 = x$, and where the data are X_{t_i}, $i = 0,\dots,297$. The conditional distribution function is easily determined by simulation. If the data have been generated by the model (2.1) with the parameter values (β,γ,σ^2), then the U_i-s are independent and uniformly distributed in the unit interval. The simulations to determine the conditional distribution function were done using the estimated parameter values. The first two plots in the second row of the Figures 4.5 – 4.8 show a histogram of the uniform residuals and a quantile plot comparing their empirical distribution to the uniform distribution. The four quantile plots do not indicate any systematic deviation from the identity line, whereas the histograms seem to have a bit too many observations near 0.5. This indication of an over-fit is probably due to the fact that the estimated parameters were used in the simulation. The graphical model diagnostics have been supplemented by formal tests of the hypothesis that the residuals are uniformly distributed. Three test statistics were calculated, the Kolmogorov-Smirnov test statistics, the χ^2 goodness-of-fit test statistics with the interval $[0,1]$ divided into ten intervals, and the test statistic $-2\sum_{i=1}^{n}\log(U_i)$ which is χ^2-distributed with $2n$ degrees of freedom if the residuals are uniform and independent. The p-values are given in Table 4.3 and do not give us reason to reject the model. To get an impression of possible deviations from independence of the residuals, the two plots in the last column of the Figures 4.5 – 4.8 show the points (i,U_i), $i = 1,\dots,297$ and the points (U_i,U_{i+1}), $i = 1,\dots,296$, respectively. These plots show no particular pattern that contradicts the independence of the residuals.

	K-S	χ^2	$-2\sum\log(U_i)$
Danish krone	48.6%	28.7%	83.1%
French franc	6.7%	4.8%	68.6%
Belgian franc	27.1%	9.4%	92.2%
Dutch guilder	29.6%	11.3%	86.9%

Table 4.3: The p-values of the tests for uniformity of residuals

In conclusion, the above investigations indicate that the Jacobi diffusion (2.1) is a reasonable model of exchange rates in a target zone between realignments of the central parity. The only serious problem is the apparent regime change for the Belgian franc, which might well reflect a change of policy.

5 An alternative model

In the previous section we concluded that the Jacobi diffusion (2.1) fits our data well. However, a close look at the Figures 4.1 – 4.4 might lead one to think that the random variation in the data is not smaller near the edge of the target zone than near the central parity, as assumed in the Jacobi model. This has lead us to study an alternative to the Jacobi diffusion with similar properties except that the diffusion coefficient is constant. The alternative model is given by the stochastic differential equation

$$dX_t = -\rho\frac{\sin\left(\frac{1}{2}\pi(X_t-m)/z\right)-\varphi}{\cos\left(\frac{1}{2}\pi(X_t-m)/z\right)}dt+\sigma dW_t, \tag{5.1}$$

where $\rho > 0$, $\varphi \in (-1, 1)$, $z > 0$, $m \in \mathbb{R}$, and where W is a standard Wiener process. The state space is $(m - z, m + z)$. The model (5.1) is a generalization of the Ornstein-Uhlenbeck process on $(-\pi/2, \pi/2)$ introduced in Kessler & Sørensen (1999), which is obtained when $\gamma = 0$, $m = 0$, and $z = \pi/2$. The economic interpretation of (5.1) is that the market volatility is constant, but that the central banks intervene very forcefully when the exchange rate comes near the boundaries to try to keep it away from them. Whether they succeed or not depends (within the model) on the strength of the intervention as we shall now see. In the following, we will consider mainly the natural form of (5.1), i.e. the case $m = 0$ and $z = \pi/2$.

The model (5.1) is very closely related to the Jacobi diffusion. To see this, suppose that X solves the stochastic differential equation (2.1) with $m = 0$ and $z = 1$, and define Y by $Y_t = \sin^{-1}(X_t)$. Then it follows by Ito's formula that

$$dY_t = -\rho \frac{\sin(Y_t) - \varphi}{\cos(Y_t)} dt + \sigma d\tilde{W}_t,$$

where $\rho = \beta - \frac{1}{2}\sigma^2$, $\varphi = \beta\gamma/(\beta - \frac{1}{2}\sigma^2)$, and where $\tilde{W}$ is the standard Wiener process given by $d\tilde{W}_t = \mathrm{sgn}(Y_t)\, dW_t$. Here $\mathrm{sgn}(x)$ denotes the sign of a real number x. To see that $\tilde{W}$ is a standard Wiener process, consider its quadratic variation.

From the connection between the alternative model and the Jacobi diffusion it follows that the solution to (5.1) is ergodic if and only if $\rho \geq \frac{1}{2}\sigma^2$ and $-1+\sigma^2/(2\rho) \leq \varphi \leq 1-\sigma^2/(2\rho)$. If (5.1) is used to model exchange rates in a target zone, ρ expresses the strength of the intervention of the central banks, φ measures the asymmetry between the two currencies, while σ^2 expresses the volatility of the market. To have a stationary situation, the strength of the intervention must be sufficiently strong compared both to the market volatility and to the asymmetry between the currencies. If the restrictions on the parameters are not satisfied, the boundaries can be reached in a finite time with positive probability.

The restrictions on the parameters that ensure ergodicity could, of course, also be obtained directly by considering the scale and speed measures of the process given by (5.1). By normalizing the speed measure, it is seen that the invariant distribution (for $m = 0$ and $z = \pi/2$) has density function

$$\pi(x; \rho, \varphi, \sigma^2) = \frac{(2\cos(x))^{2\rho/\sigma^2}}{B\left(\rho(1+\varphi)/\sigma^2 + \frac{1}{2}, \rho(1-\varphi)/\sigma^2 + \frac{1}{2}\right)} \left(\frac{1+\sin(x)}{1-\sin(x)}\right)^{\rho\varphi/\sigma^2}, \qquad (5.2)$$

$-\pi/2 < x < \pi/2$, where as earlier B denotes the beta-function. Note that $\pi(-x; \rho, \varphi, \sigma^2) = \pi(x; \rho, -\varphi, \sigma^2)$. The distribution given by (5.2) is of course the transformation of (2.2) by $\sin^{-1}$.

The fact that the model (5.1) can be obtained from the Jacobi diffusion by the transformation $\sin^{-1}$ makes it very likely, that (5.1) will fit the data at least as well as the Jacobi diffusion. In the centre of the state space, the transformation $\sin^{-1}$ is not far from the identity so that the two models are very similar in the part of the state space where most of the data are. Near the boundaries the transformation is non-linear in a way that causes the diffusion coefficient of the transformed process to be constant, which is exactly what is needed to get a better fit near the boundaries. It would be interesting to investigate in detail how well the alternative model fits the interest rate data, but this is beyond the scope of the present paper.

Let us end this section by noting that estimation for the alternative model is exactly as easy as for the Jacobi diffusion. This is not surprising given the connection between the two models. From this connection it follows that the eigenfunctions of the generator of (5.1) are

$$\varphi_n(x; \rho, \varphi, \sigma, m, z) = P_n^{(\rho(1-\varphi)\sigma^{-2}-\frac{1}{2},\, \rho(1+\varphi)\sigma^{-2}-\frac{1}{2})}\left(\sin(\tfrac{1}{2}\pi x/z - m)\right), \qquad (5.3)$$

where $P_n^{(a,b)}(x)$ denotes the Jacobi polynomial of order n, see (2.4). The eigenvalue corresponding to φ_n is

$$\lambda_n(\rho,\varphi,\sigma) = n\left(\rho + \tfrac{1}{2}n\sigma^2\right). \tag{5.4}$$

So again it is natural to use an estimating function based on the eigenfunctions, which will also here have the advantages discussed in Section 2. In particular, the eigenfunctions have a form that allows explicit calculation of the optimal estimating function as explained in the Appendix.

6 Option pricing

A main reason for developing continuous time financial models and related statistical methods is to be able to calculate the price of derivative assets by the elegant no-arbitrage theory of continuous trading. Classical references are Black & Scholes (1973) and Harrison & Pliska (1981). Here we will discuss pricing of options on the exchange rate. The discussion follows the line of arguments in De Jong, Drost & Werker (2001), so we will not go into details.

Let r_t and r_t^f denote the instantaneous domestic and foreign interest rates, respectively, and let S_t be the exchange rate between the domestic and the foreign currency. Suppose we have a foreign money market account with value V_t^f in the foreign currency and value V_t in the domestic currency. Then $dV_t^f = r_t^f V_t^f dt$ and $V_t = S_t V_t^f$. Finally let $D_t = \exp\left(-\int_0^t r_s ds\right) S_t V_t^f$ be the discounted value of the account in the domestic currency. If $X_t = \log(S_t)$ is modelled by (2.1), then by Ito's formula

$$dD_t = \left\{r_t^f - r_t - \beta[X_t - (m+\gamma z)] + \tfrac{1}{2}\sigma^2\left[z^2 - (X_t - m)^2\right]\right\} D_t dt + \sigma D_t \sqrt{z^2 - (X_t - m)^2} dW_t.$$

So far we have not made any assumptions about the variation of the interest rates. It seems a reasonable assumption that the difference $r_t^f - r_t$ depends on X_t in such a way that there are no arbitrage possibilities. This is equivalent to assuming that there exists an equivalent probability measure under which D is a martingale, see Harrison & Pliska (1981). We shall discuss explicit specifications of $r_t^f - r_t$ below. There exists a stochastic process $\tilde{W}$, which is a standard Wiener process under the equivalent martingale measure, such that

$$dD_t = \sigma D_t \sqrt{z^2 - (X_t - m)^2} d\tilde{W}_t,$$

also under this measure. Since $S_t = V_t / V_t^f$, it follows from Ito's formula that under the risk-neutral measure

$$dX_t = \left\{r_t - r_t^f - \tfrac{1}{2}\sigma^2\left[z^2 - (X_t - m)^2\right]\right\} dt + \sigma\sqrt{z^2 - (X_t - m)^2} d\tilde{W}_t \tag{6.1}$$

and

$$dS_t = (r_t - r_t^f) S_t dt + \sigma S_t \sqrt{z^2 - (X_t - m)^2} d\tilde{W}_t. \tag{6.2}$$

Thus the no-arbitrage assumption implies that the "uncovered interest rate parity" holds under the martingale measure, as also pointed out by Christensen, Lando & Miltersen (1998).

Not every specification of how $r_t^f - r_t$ depends on X_t is in accordance with the assumption of no arbitrage. For instance, if $r_t^f - r_t$ equals a constant ν, it is obvious that the upper or lower (dependent on the sign of ν) boundary of the state space can be hit by X_t in finite time with positive probability. This claim can be substantiated by a simple calculation involving the scale measure. The probability measure on the canonical space corresponding to (6.1) can therefore not be equivalent to the original measure. Also a simple economic consideration makes it clear that there would be arbitrage opportunities if the difference between the domestic and foreign interest rates did not depend on the exchange rate. To see what problems might occur, consider the simple specification

$$r_t^f - r_t = \beta_1(X_t - m_1), \tag{6.3}$$

where $\beta_1 > 0$ and $m_1 \in (m-z, m+z)$. The economic interpretation is that the difference between the two interest rates is proportional to the deviation of the exchange rate from a level that the market considers to be reasonable. With the specification (6.3), X_t solves the equation

$$dX_t = \left\{-\beta_1(X_t - m_1) - \tfrac{1}{2}\sigma^2\left[z^2 - (X_t - m)^2\right]\right\}dt + \sigma\sqrt{z^2 - (X_t - m)^2}\,d\tilde{W}_t \tag{6.4}$$

under the risk-neutral measure. In analogy with the discussion in Section 2, the solution to (6.4) is ergodic if and only if $\beta_1 \geq \sigma^2$ and $-1 + \sigma^2/\beta_1 \leq (m_1 - m)/z \leq 1 - \sigma^2/\beta_1$, as follows from an inspection of the behaviour of the scale measure at the boundaries. If these restrictions on β_1 and m_1 are not satisfied, the solution to (6.4) can hit one of the boundaries of the state space in finite time with positive probability. Again this implies that the probability measure corresponding to (6.4) cannot be equivalent to the original measure, so there are arbitrage opportunities. We see that there must be a suitable balance between how far m_1 is from the central parity, the constant of proportionality β_1, and the market volatility. Obviously other specifications of $r_t^f - r_t$ are possible, but it should always be ensured that the solution to (6.1) cannot hit the boundary in finite time.

A similar discussion can be carried through if $X_t = \log(S_t)$ is modelled by (5.1). If the difference $r_t^f - r_t$ is specified in such a way that there is no arbitrage opportunities, then under the risk-neutral measure

$$dX_t = \left\{r_t - r_t^f - \tfrac{1}{2}\sigma^2\right\}dt + \sigma d\tilde{W}_t \tag{6.5}$$

and

$$dS_t = (r_t - r_t^f)S_t dt + \sigma S_t d\tilde{W}_t. \tag{6.6}$$

In this case a constant difference between r_t^f and r_t would imply that X is a Brownian motion with drift under the risk-neutral measure, so obviously equivalence does not hold. Also the linear specification (6.3) goes wrong, as it implies that X is the Ornstein-Uhlenbeck process on $\mathbb{R}$. The difference between the two interest rates must here necessarily go to infinity as X_t goes to one of the boundaries to ensure equivalence and hence no-arbitrage. This might seem extreme, but central banks have been seen to react in ways reminiscent of this when speculation has pushed the exchange rate close to the boundary of a target zone in a situation where a readjustment of the central parity was politically unacceptable. An example is the interest rate of 500 per cent in Sweden in the days September 17 – 20 1992. A way to avoid that the difference between the two interest rates must go to infinity is discussed in the next section. One possible specification of $r_t^f - r_t$ that fits well with the alternative model (5.1) is

$$r_t^f - r_t = \rho_1 \frac{\sin\left(\frac{1}{2}\pi(X_t - m)/z\right) - \varphi_1}{\cos\left(\frac{1}{2}\pi(X_t - m)/z\right)}, \tag{6.7}$$

where $\rho_1 > 0$ and $\varphi_1 \in (-1, 1)$. The interpretation is similar to that of (6.3). The difference $r_t^f - r_t$ depends on how far X_t is from a certain level $m_1 = m + 2z\sin^{-1}(\varphi_1)/\pi$, but here the dependence is non-linear, in particular near the boundaries. With the specification (6.7), X_t solves the equation

$$dX_t = \left\{ -\rho_1 \frac{\sin\left(\frac{1}{2}\pi(X_t - m)/z\right) - \varphi_1}{\cos\left(\frac{1}{2}\pi(X_t - m)/z\right)} - \tfrac{1}{2}\sigma^2 \right\} dt + \sigma d\tilde{W}_t. \tag{6.8}$$

By considering the scale measure, we see that the solution to (6.8) is ergodic and does not hit the boundaries if and only if $\rho_1 \geq \frac{1}{2}\sigma^2$ and $-1 + \sigma^2/(2\rho_1) \leq \varphi_1 \leq 1 - \sigma^2/(2\rho_1)$. As with the model (2.1), there are arbitrage opportunities if there is not a suitable balance between how far m_1 is from the central parity, the constant of proportionality ρ_1, and the market volatility. Other specifications $r_t^f - r_t = \Psi(X_t)$ are possible, of course, but it is necessary that $\Psi(x)$ tends to infinity at the boundary $m + z$ and to minus infinity at $m - z$ sufficiently fast that the scale measure diverges at both boundaries.

Once a model of X_1 and $r_t^f - r_t$ has been chosen, the price of a derivative asset, where the pay-off, Q, is some functional of X, can be found by calculating the expectation of the discounted value of Q when X is given by (6.1) or (6.5). This can typically be easily done by simulating X, for instance with one the methods in Kloeden & Platen (1999). As an example consider a European call option on the exchange rate with maturity T and strike price K. The domestic option price is

$$\tilde{E}\left(\exp\left(-\int_0^T r_s ds\right) \max\{S_T - K, 0\}\right),$$

where $\tilde{E}$ denotes the expectation under the risk-neutral measure, i.e. when S is given by (6.2) or (6.6).

As expected the price of a contingent claim on the exchange rate does not depend on the drift of the diffusion under the "physical" probability measure. Let us therefore conclude this section by listing two reasons why it is important to model the drift carefully. One reason is that to obtain good estimates of the parameters in the diffusion coefficient, it is necessary to have a realistic model of the drift and good estimators for the drift parameters. Otherwise it would be impossible to determine with any precision what part of the random variation in the data is due to the diffusion and what part is due to the drift. An estimator of the quadratic variation is usually too crude, except when the sampling frequency is very high. A second reason is that it is impossible to determine whether the martingale measure is equivalent or not if the physical measure has not been specified. As we have seen, this can be a real problem.

7 Realignments of the central parity and two-factor models

A problem that has so far been ignored in this paper is that the central parity of the target zone can occasionally be changed to better fit the actual relative value of the two currencies when it becomes obviously unreasonable. Also the width of the target zone can be changed, but this happens much more rarely.

A simple, but rather unrealistic model of realignments of the central parity was proposed by De Jong, Drost & Werker (2001) who, following other authors, modelled the realignment times as a Poisson process with constant intensity and equaled the new central parity to the exchange rate at the realignment time. In this way a simple diffusion with jumps is obtained.

It would be straightforward to similarly build realignments into our two models, but given the lack of realism of a Poisson model, we did not find it worthwhile to do so.

Empirical evidence suggests that most realignments occur when the exchange rate is close to one of the boundaries of the target zone, see Ball & Roma (1993). It would not be difficult to propose other, and perhaps more realistic, specifications of the intensity λ_t of the counting process of the realignment times. Obvious possibilities are $\lambda_t = \nu \exp(\xi |X_{t-} - m|)$ or $\lambda_t = \nu \int_{\tau_{i-1}}^{t} \exp(\xi_1 |X_s - m| - \xi_2(t-s))ds$, when $\tau_{i-1} < t \leq \tau_i$. Here the τ_i-s are the realignment times, and the parameters ν, ξ, ξ_1, and ξ_2 are non-negative. In the first case a diffusion with jumps is obtained, in the latter a non-Markovian model is obtained. Estimation of parameters in the intensity λ_t is not difficult because the jump-times are observed. Statistical methods have been available for many years, see e.g. Jacobsen (1982). It should be noted that we can avoid the conclusion discussed in the previous section that the no-arbitrage assumption implies that under the alternative model (5.1) the difference between the domestic and the foreign interest rates must go to infinity when the exchange rate is near one of the boundaries of the target zone by assuming that the realignment intensity λ_t goes to infinity in a suitable way when X_t is near the boundaries. This is because when the risk of realignment is included in the model, a term equal to the product of the realignment intensity and the jump size appears in the drift of X_t under the martingale measure, see Christensen, Lando & Miltersen (1998).

A more realistic model should probably involve a second variable, such as the shadow exchange rate used in the target zone models of Christensen, Lando & Miltersen (1998) and Rangvid & Sørensen (2001). These authors include in their model a second diffusion process Y, which represents the free floating exchange rate in the absence of a target zone (given by the relative purchasing power parity assumption). Christensen, Lando & Miltersen (1998) specified the intensity of the counting process of the realignment times as

$$\lambda_t = \nu_1 \max\left\{0, \frac{(S_{t-} - \mu_{t-})}{\mu_{t-}} \cdot \frac{(Y_{t-} - S_{t-})}{S_{t-}}\right\} + \nu_2, \tag{7.1}$$

where μ_t denotes the central parity, which now depends on time, and $\nu_1, \nu_2 \geq 0$. With this specification, the probability of a realignment is increased both because of speculative pressure when the exchange rate is far from the central parity, and because of the imbalance between the exchange rate and fundamental macroeconomic conditions when the exchange rate is far from the shadow exchange rate. However, if the exchange rate is not between the central parity and the shadow exchange rate, a realignment cannot occur, which makes good economic sense. Rangvid & Sørensen (2001) specified λ_t in a slightly different way. If we wish to supplement one of our two models with a shadow exchange rate process Y, this process should be correlated with the exchange rate process because the two processes are driven by the same fundamental economic events. If we model this simply by assuming that the two Wiener processes that drive the two processes Y and X are correlated (or even identical), then the results about X that we have derived in this paper are still valid. In particular, the estimation method discussed is still applicable. However, the parameters appearing in (7.1) and in the specification of Y cannot be estimated from exchange rate data because Y cannot be observed. If the drift under the martingale measure were known, we could use the fact that the uncovered interest rate parity holds under the martingale measure to calculate Y from the exchange rate and the interest rate difference $r_t^f - r_t$, but since the model with realignment risk is not complete, we do not know the drift under the martingale measure. A way to estimate the remaining parameters would be to supplement the exchange rate data with data on currency option prices. Rangvid & Sørensen (2001) assumed that the uncovered interest rate parity holds under the physical probability measure too and used this assumption to calculate Y. They then estimated the parameters in a full model of X and Y by approximate maximum likelihood estimation based on an Euler approximation, a method that is known to be biased.

A different approach would be to let the drift of X depend on the process Y. For the model (2.1), a natural way to do this would be to let the asymmetry parameter γ depend on Y, e.g.

$$\gamma_t = \left(\frac{Y_t}{\mu_t}\right)^\rho / \left(1 + \left(\frac{Y_t}{\mu_t}\right)^\rho\right),$$

where $\rho \geq 0$. For the model (5.1), the parameter φ could be assumed to depend on Y_t in the same way. These models would be particular cases of the general setup in Christensen, Lando & Miltersen (1998), but different from the specification in their worked example. Thus their option pricing method would apply. Unfortunately, most of the results presented in this paper would not hold for such a model, but because the dynamics of X depends more directly on the state of Y, it would probably be possible to estimate all the parameters of the model from exchange rate data by means of one of the statistical methods that are generally applicable to two-factor models such as stochastic volatility models. Examples are the method of prediction-based estimating functions, see Sørensen (2000), Markov chain Monte Carlo methods, or one of the indirect inference methods. However, the inclusion of a shadow exchange rate in the model increases the necessary computational effort substantially, both when estimating the parameters of the model and when calculating prices of contingent claims, and it should be carefully considered whether this is justified by the implied increase in the precision of the calculated option prices.

Option pricing is more difficult if the risk of realignments of the central parity is taken into account, because then the model describes an incomplete market so that derivative assets cannot be priced solely by a no-arbitrage argument (unless a further risky financial instrument is added). Some decision must be made about how to treat the realignment risk. De Jong, Drost & Werker (2001) chose not to put a risk-premium on the realignment risk. More specifically, among all the possible martingale measures they chose the one that changes the diffusion in a way similar to what was described in the previous section, while it does not change the distribution of the realignments. In fact, there are infinitely many martingale measures under which the intensity of the counting process of the realignment times is changed in infinitely many ways. Christensen, Lando & Miltersen (1998) also did not discuss how to choose the martingale measure. Instead they completed their model by adding two risky assets, e.g. two currency options with different maturities, and discussed how to hedge the option by positions in a domestic money market account, the foreign currency and the two options.

Acknowledgements

We are grateful to F. C. Drost for putting the data at our disposal and to C. Sørensen for helpful comments on an earlier version of the paper. Most of Michael Sørensen's contribution to this paper was done while visiting the Department of Statistics and the Financial Mathematics Program at the University of Chicago. The research of Michael Sørensen was supported by the Centre for Analytical Finance and the Danish Mathematical Finance Network, both financed by the Danish Social Science Research Council, by the European Commission through the Research Training Network DYNSTOCH under the Human Potential Programme, and by MaPhySto – A Network in Mathematical Physics and Stochastics funded by The Danish National Research Foundation.

Appendix: Optimal estimation

In this appendix we shall demonstrate that for diffusions models where the eigenfunctions are polynomials of a particular fixed function, the optimal estimating function based on eigenfunctions can be found explicitly. Both of the models considered in the present paper have eigenfunctions of this type.

Consider a diffusion model indexed by a p-dimensional parameter θ for which the eigenfunctions are of the form

$$\varphi_i(x;\theta) = \Pi_i(\kappa(x);\theta),$$

where Π_i is a polynomial of degree i

$$\Pi_i(y;\theta) = \sum_{j=0}^{i} a_{i,j}(\theta)\, y^j,$$

and κ is a real function, independent of θ, defined on the state space of the diffusion. Denote the corresponding eigenvalues by $\lambda_i(\theta)$ and the state space by (ℓ, r). Assume that we want to draw inference about θ from the data $X_{t_0}, X_{t_1}, \ldots, X_{t_n}$ ($t_0 = 0$). The optimal estimating function based on the first N eigenfunctions is (see Kessler & Sørensen (1999))

$$G_n^*(\theta) = \sum_{i=1}^{n} B(\Delta_i, X_{t_{i-1}};\theta) C(\Delta_i, X_{t_{i-1}};\theta)^{-1} h(\Delta_i, X_{t_{i-1}}, X_{t_i};\theta),$$

where $\Delta_i = t_i - t_{i-1}$, $h = (h_1, \ldots, h_N)^T$ with $h_j(\Delta, x, y;\theta)$, $j = 1, \ldots, N$ given by (2.7), $B(\Delta, x;\theta) = \{b_{ij}(\Delta, x, \theta)\}$ is the $p \times N$-matrix with entries

$$b_{ij}(\Delta, x, \theta) = \sum_{k=0}^{j} \partial_{\theta_i} a_{j,k}(\theta) \int_\ell^r \kappa(y)^k p(\Delta, x, y;\theta)dy - \partial_{\theta_i}(e^{-\lambda_j(\theta)\Delta}\varphi_j)(x;\theta)$$

and $C(\Delta, x;\theta) = \{c_{ij}(\Delta, x, \theta)\}$ is the $N \times N$-matrix with entries

$$c_{i,j}(\Delta, x, \theta) = \sum_{r=0}^{i}\sum_{s=0}^{j} a_{i,r}(\theta) a_{j,s}(\theta) \int_\ell^r \kappa(y)^{r+s} p(\Delta, x, y;\theta)dy - e^{-[\lambda_i(\theta)+\lambda_j(\theta)]\Delta}\, \varphi_i(x;\theta)\varphi_j(x;\theta).$$

Here $y \mapsto p(\Delta, x, y;\theta)$ denotes the transition density, i.e. the conditional density of X_Δ given that $X_0 = x$. Optimality is in the sense of Godambe & Heyde (1987), see also Heyde (1997).

In the expressions for the matrices B and C only the integrals $\int_\ell^r \kappa(y)^k p(\Delta, x, y;\theta)dy$, $k = 1, \ldots, 2N$ are not explicit. These integral can, however, be found by integrating both sides of the expression for the eigenfunctions $\varphi_i(x;\theta) = \Pi_i(\kappa(x);\theta)$ with respect to $p(\Delta, x, y;\theta)$ for $i = 1, \ldots, 2N$. From the fact that the functions φ_i are eigenfunctions it follows that

$$e^{-\lambda_i(\theta)}\varphi_i(x;\theta) = \sum_{j=0}^{i} a_{i,j}(\theta) \int_\ell^r \kappa(y)^j p(\Delta, x, y;\theta)dy$$

for $i = 1, \ldots, 2N$. From these linear equations we can obtain the integrals, and thus an explicit expression for the optimal estimating function based on the first N eigenfunctions.

Under regularity conditions like those in Kessler & Sørensen (1999), the estimator $\hat\theta_n$ obtained by solving the estimating equation $G_n^*(\theta) = 0$ exists with a probability that goes to one as $n \to \infty$, is consistent, and is asymptotically normal. Specifically, if θ_0 denotes the true value of θ and the sampling times are equidistant with $t_i - t_{i-1} = \Delta$, then

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{D}} N\left(0\,,\, I(\theta_0)^{-1}\right)$$

as $n \to \infty$, where

$$I(\theta) = \int_\ell^r B(\Delta, x; \theta) C(\Delta, x; \theta)^{-1} B(\Delta, x; \theta) \pi(x; \theta) dx$$

with $\pi(x;\theta)$ denoting the density of the invariant distribution of the diffusion model.

References

Abramowitz, M. & Stegun, I. A. (1964). *Handbook of Mathematical Functions.* National Bureau of Standards, Washington D.C.

Ball, C. A. & Roma, A. (1993). "A jump diffusion model for the European Monetary System". *Journal of International Money and Finance*, 12:475–492.

Bibby, B. M.; Skovgaard, I. M. & Sørensen, M. (2003). "Diffusion-type models with given marginal and autocorrelation function". Preprint 2003–5, Department of Applied Mathematics and Statistics, University of Copenhagen.

Black, F. & Scholes, M. (1973). "The pricing of options and corporate liabilities". *Journal of Political Economy*, 81:637–659.

Christensen, P. O.; Lando, D. & Miltersen, K. R. (1998). "State dependent realignments in target zone currency regimes". *Review of Derivatives Research*, 1:295–323.

De Jong, F.; Drost, F. C. & Werker, B. J. M. (2001). "A jump-diffusion model for exchange rates in a target zone". *Statistica Neerlandica*, 55:270–300.

Godambe, V. P. & Heyde, C. C. (1987). "Quasi likelihood and optimal estimation". *International Statistical Review*, 55:231–244.

Gourieroux, C. & Jasiak, J. (2003). "Jacobi process and smooth transitions". Working paper, CREST.

Harrison, M. & Pliska, S. (1981). "Martingales and stochastic integrals in the theory of continuous trading". *Stoch. Proc. Appl.*, 20:381–408.

Heyde, C. C. (1997). *Quasi-Likelihood and Its Application.* Springer-Verlag, New York.

Jacobsen, M. (1982). *Statistical Analysis of Counting Processes.* Springer-Verlag, New York. Lecture Notes in Statistics 12.

Kessler, M. & Sørensen, M. (1999). "Estimating equations based on eigenfunctions for a discretely observed diffusion process". *Bernoulli*, 5:299–314.

Kloeden, P. E. & Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations.* 3rd revised printing. Springer-Verlag, New York.

Krugman, P. (1991). "Target zones and exchange rate dynamics". *Quarterly Journal of Economics*, 106:669–682.

Pedersen, A. R. (1994). "Uniform residuals for discretely observed diffusion processes". Research Report No. 292, Department of Theoretical Statistics, University of Aarhus.

Press, W. H.; Teukolsky, S. A.; Vetterling, W. T. & Flannery, B. P. (1992). *Numerical Recipes in C*. Cambridge University Press, Cambridge.

Rangvid, J. & Sørensen, C. (2001). "Determinants of the implied shadow exchange rates from a target zone". *European Economic Review*, 45:1665–1696.

Sørensen, M. (2000). "Prediction-based estimating functions". *Econometrics Journal*, 3:123–147.