Kiel
Working Papers

Kiel Institute
for the World Economy

Title: Inference for Systems of Stochastic Differential Equations from Discretely Sampled data: A Numerical Maximum Likelihood Approach

by Prof. Dr. Thomas Lux

No. 1781| July 2012

Web: www.ifw-kiel.de

Kiel Working Paper No. 1781| Month Year

Title* Inference for Systems of Stochastic Differential Equations from Discretely Sampled data: A Numerical maximum Likelihood Approach

Author: Prof. .Dr. Thomas Lux

Abstract:
Maximum likelihood estimation of discretely observed diffusion processes is mostly hampered by the lack of a closed form solution of the transient density. It has recently been argued that a most generic remedy to this problem is the numerical solution of the pertinent Fokker-Planck (FP) or forward Kolmogorov equation. Here we expand extant work on univariate diffusions to higher dimensions. We find that in the bivariate and trivariate cases, a numerical solution of the FP equation via alternating direction finite difference schemes yields results surprisingly close to exact maximum likelihood in a number of test cases. After providing evidence for the efficiency of such a numerical approach, we illustrate its application for the estimation of a joint system of short-run and medium run investor sentiment and asset price dynamics using German stock market data.

Keywords: Stochastic differential equations, numerical maximum likelihood, Fokker-Planck equation, finite difference schemes, asset pricing

JEL classification: C58, G12, C13

Prof. Dr. Thomas Lux
Kiel Institute for the World Economy
24100 Kiel, Germany
Telephone: +49 -431-8814-278
E-mail: Thomas.lux@ifw-kiel.de

Bank of Spain
Chair of Computational Economics
Department of Economics, University Jaume I
Castellón, Spain

Inference for Systems of Stochastic Differential Equations from Discretely Sampled data: A Numerical Maximum Likelihood Approach

Thomas Lux*

July 9, 2012

Department of Economics, University of Kiel,
Kiel Institute for the World Economy
Kiel, Germany
and
Bank of Spain Chair of Computational Economics
Department of Economics, University Jaume I
Castellón, Spain

Abstract: Maximum likelihood estimation of discretely observed diffusion processes is mostly hampered by the lack of a closed form solution of the transient density. It has recently been argued that a most generic remedy to this problem is the numerical solution of the pertinent Fokker-Planck (FP) or forward Kolmogorov equation. Here we expand extant work on univariate diffusions to higher dimensions. We find that in the bivariate and trivariate cases, a numerical solution of the FP equation via alternating direction finite difference schemes yields results surprisingly close to exact maximum likelihood in a number of test cases. After providing evidence for the efficiency of such a numerical approach, we illustrate its application for the estimation of a joint system of short-run and medium run investor sentiment and asset price dynamics using German stock market data.

Keywords: *Stochastic differential equations, numerical maximum likelihood, Fokker-Planck equation, finite difference schemes, asset pricing.*

JEL classification: C58, G12, C13

*Helpful comments by François Bouchut and participants of a Workshop on "Heterogeneity and Networks in Financial Markets" at GREQAM Aix-en-Provence as well as the careful reading and constructive comments by an anonymous reviewer are gratefully acknowledged.
Contact adress: Thomas Lux, Department of Economics, University of Kiel, Olshausen Str. 40, 24118 Kiel, Germany, E-Mail: lux@bwl.uni-kiel.de

1 Introduction

Stochastic differential equations (SDEs) are frequently used in such diverse fields as physics, biology and financial economics. Typical examples of applications of SDEs include many areas of statistical mechanics, population dynamics as well as interest rate models or derivative pricing. In all of these areas, estimation of the parameters of SDEs from discretely sampled data is a natural concern. Indeed, a plethora of estimation methods exists in the literature that approach the problem of statistical inference from quite diverse angles. Given its asymptotic efficiency and consistency, exact maximum likelihood (i.e., maximum likelihood estimation based on a closed-form solution of the transient density of the state variables) would be the method of choice where applicable. Unfortunately, the applicability of exact maximum likelihood (EML) is mostly (except for the most elementary linear cases) hampered by the absence of a closed-form solution for the transient density.

With this unfortunate situation, a number of alternatives can be used. Hurn *et al.* (2009) in their survey of estimation methods for single variable, time-homogenous SDEs, distinguish between two classes of methods: Likelihood- based approaches and what they call "sample DNA matching procedures". The latter group is characterized by its attempt to match certain characteristics of the data, e.g. moment conditions, the shape of the characteristic function or that of the marginal density. Likelihood-based methods of inference, in contrast, take the likelihood function as their starting point and are numerical or simulation-based approaches for approximate maximization of the likelihood. In simulated ML, the conditional density is approximated via Monte Carlo simulation. More refined approaches in this area use importance sampling and Markov Chain Monte Carlo techniques. As for numerical approximations of the likelihood function, the best-known approaches are "discrete" ML using an Euler approximation and Hermite polynomial expansions (Ait-Sahalia, 1999).

Unfortunately, the applicability of these methods is restricted by some serious disadvantages: "discrete" ML comes with inconsistent parameter estimates, while Hermite expansions are not always easy to apply. Hurn *et al.* (2009) note that, in principle, a more generic approach would consist in a numerical approximation to the transitional density. Since the transitional density of SDEs is known to obey the so-called Fokker-Planck or forward Kolmogorov equation, a convenient strategy in the absence of a closed-form solution would be numerical approximation of the dynamics imposed by the Fokker-Planck equation (FPE). Since the

FPE is a partial differential equation, various numerical approximation schemes could be used. Given the accuracy of an approximation method, one might infer asymptotic efficiency and consistency properties of the numerical likelihood maximization (Poulsen, 1999). Hurn *et al.* (2010) use a Crank-Nicolson finite difference scheme as well as a related scheme adopted for the cumulative distribution function and demonstrate that in Monte Carlo simulations its performance is among the best of a large variety of estimation methods. Besides efficiency, the main advantage of this approach is that it is the most generic method since its basic building blocks can be encoded once for all and used for most families of SDEs in exactly the same way. In contrast, many of the other methods require a problem-specific design of the algorithm. Surprisingly, while numerical approximations of partial differential equation are quite common in many areas, using a numerical approximation of the FPE in order to perform maximum likelihood estimation is a relatively new development. Even recent textbooks such as Iacus (2008) do not mention this possibility. While Poulsen (1999) already discussed this approach, the only subsequent applications to my knowledge are Jensen and Poulsen (2002) and Hurn *et al.* (2009, 2010). Closely related, yet somewhat different in their scope are two papers by Lux (2009, 2012) whose aim is to estimate the parameters of a dynamic model of agents' interaction in a finance or economics setting. While this approach is not based on phenomenological SDEs, the more complicated jump Markov process for an ensemble of agents can also be approximated by a Fokker-Planck equation. Although the FPE is only an approximate law of motion in this case (while it is exact for SDEs), statistical inference proceeds along the same lines as outlined above. Lux (2012) estimates the parameters of a dynamic process with two or three state variables which requires the numerical approximation of a FPE in higher dimensions.

For SDEs, all available literature is restricted to the one-dimensional case. The purpose of this paper is to go beyond univariate SDEs. I first introduce a general framework for a system of two SDEs that covers a number of interesting cases like the Ornstein-Uhlenbeck or Vasicek process, the double-well potential (Iacus, 2008, c. 1.13.8) or the stochastic cusp catastrophe model. The latter types have highly nonlinear drift and have been occasionally used in economics to capture multiple equilibria (e.g. Creedy et al., 1996; Rheinlaender and Steinkamp, 2004; Barunik and Vosvrda, 2009). However, estimation of such models has so far been based on the unconditional density which presumes a fast convergence towards some equilibrium. In contrast, our approach is based on the transient density and does not rely on any assumption of a system in equilibrium. The plan of the subsequent sections is as follows: Section 2 states our problem and section

3 reviews some popular finite difference schemes that could be applied for the numerical solution of the FPE in the bivariate and trivariate case. As it turns out, correlation of innovations (i.e. the mixed derivative term of the diffusion function) is somewhat problematic and requires particular attention. Section 4 reports results on Monte Carlo simulations with various finite difference schemes in the bivariate case for two correlated Brownian motions and various combinations of processes with linear and nonlinear drift functions. Section 5 extends our Monte Carlo analysis to the trivariate case. Using the best performing FD algorithm we finally proceed to an empirical application in section 6 investigating the joint dynamics of two sentiment indices and the stock price index DAX of the German share market. Section 7 concludes.

2 Problem Formulation

Our parameter estimation problem can be concretized as follows. Available is an empirical sample of $T+1$ observations $X_0, ..., X_T$ at times $t_0, \ldots, t_T$ where each observation X_t is multi-variate with n different variables, i.e. our sample consists of $\{X_t\}_{t=0}^{T} = \{x_{1,t}, x_{2,t}, \ldots, x_{n,t}\}_{t=0}^{T}$. Our hypothesized data-generating process for this discrete multivariate sample is an n-dimensional time-homogenous system of stochastic differential equation:

$$dX = \mu(X;\theta)dt + \sqrt{g(X;\theta)}dW. \tag{1}$$

In eq.(1), $\mu(X;\theta)$ is the drift and $g(X;\theta)$ the diffusion of our system of equations, θ is a vector of unknown parameters and W is an n-dimensional Wiener process with components $W^1, W^2, \ldots, W^n$ that are independent scalar Wiener processes. The maximum likelihood estimates of θ are obtained by minimizing the negative log-likelihood function:

$$-log\mathcal{L}(\theta) = -logf_0(X_0|\theta) - \sum_{S=0}^{T-1} logf(X_{S+1}|X_S;\theta), \tag{2}$$

where $f_0(X_0|\theta)$ is the density of the initial state and the remaining terms $f(X_{S+1}|X_S;\theta)$ are the transient densities obtained for each set of iterations at time $S+1$ conditional on the previous set of observations at time S.

When the drift and diffusion processes are sufficiently regular functions, the transient density satisfies the so-called Kolmogorov forward or Fokker-Planck

equation (cf. Kloeden and Platen, 1992, c.4):

$$\frac{\partial f}{\partial t} = -\sum_{i=1}^{n} \partial_i[\mu_i(X;\theta)f] + \frac{1}{2}\sum_{i,j=1}^{n} \partial_i\partial_j[g_{i,j}(X;\theta)f] \tag{3}$$

with suitable initial conditions. Throughout this paper, the drift and diffusion terms will satisfy standard regularity conditions so that the existence of the Fokker-Planck equation and the uniqueness of its solution can be taken for granted. In eq. (3), $\mu_i(.)$ denotes the scalar component of the drift function associated with variable x_i $(i = 1, \ldots, n)$ and $g_{i,j}(X;\theta)$ denotes the scalar entry of the matrix of diffusion coefficients at position (i, j). In the current paper, we will focus on various forms of drift functions, but for the diffusion assume independence of X, i.e. a matrix g of constant coefficients.

If we could obtain a closed-form solution of eq.(3), exact maximum likelihood would be feasible. However, except for very simple cases, closed-form solutions are typically not known. Various numerical and simulation-based approaches can be used to obtain quasi-maximum likelihood estimates in this case. As pointed out by Hurn *et al.* (2009), numerical solution of the Fokker-Planck equation appears to be the most generic approach although it has been introduced in the literature only recently. Indeed, if a certain accuracy of the approximation can be guaranteed, this approach should be able to overcome the problem of inconsistency of simple approximations to the transient density such as the Euler-Maruyama algorithm.

3 Some Popular Finite Difference Schemes in the Bivariate Case

Numerical estimation based on a finite difference approximation to the Fokker-Planck equation in one dimension is considered in Poulsen (1999), Jensen and Poulsen (2002), Hurn *et al.* (2009) and Lux (2009). Following many other applications in various fields, the finite difference scheme chosen by most authors is the Crank-Nicolson algorithm. The reason for the almost unanimous preference is that Crank-Nicolson is unconditionally stable (a necessary requirement in the presence of unknown parameter values) and it is more accurate than many other methods (e.g., a fully implicit scheme that also is unconditionally stable). It is indeed well-known among applied researchers that Crank-Nicolson is very reliable and hard to beat even with refined methods for numerical solution of partial

differential equations in the univariate case.

The picture, however, becomes more varied when moving from the univariate case to higher dimensions. In the bivariate case, a large portfolio of finite difference schemes exists without such a clear preference as for the Crank-Nicolson scheme in the univariate case.

Some of the most versatile schemes are known as alternating direction implicit (ADI) schemes (cf. Marchuk, 1990; Ames, 1992; Morton and Mayers, 1994, and Strikwerda, 2004). Some variants of ADI schemes will be introduced below. As their defining characteristic, these approaches perform first a half (or auxiliary) step into one of two space dimensions followed by a second half (or auxiliary) step in the other direction. Using an implicit finite difference scheme in both half-steps, these algorithms boil down to computationally convenient tri-diagonal systems of equations. A very similar approach has been developed by various Russian authors under the name of splitting schemes (cf. Yanenko *et al.*, 1971, for a monograph covering most of this literature). Out of the large variety of ADI and splitting schemes we select three of the more popular examples and explore their applicability within our estimation context.

To set the stage, we start with a simple bivariate Wiener process to explore the performance of various FD schemes within a numerical ML estimation exercise. This amounts to a specification of eqs.(1) and (3) with $n = 2$, vanishing drift component $\mu(X;\theta) = 0$, and a constant variance-covariance matrix g. The vector of parameters to be estimated, then, consists of the three parameters governing g, i.e. $\theta = \{\sigma_1, \sigma_2, \rho\}$ with σ_1 and σ_2 the instantaneous diffusion of the two Wiener processes and ρ their correlation. The pertinent parabolic Fokker-Planck equation can be written as:

$$\begin{aligned} \frac{\partial f}{\partial t} &= (b_{11}\frac{\partial^2}{\partial x_1^2} + b_{12}\frac{\partial^2}{\partial x_1 \partial x_2} + b_{22}\frac{\partial^2}{\partial x_2^2})f \\ \text{with} \qquad b_{11} &= \frac{1}{2}\sigma_1^2, \qquad b_{12} = \rho\sigma_1\sigma_2, \qquad b_{22} = \frac{1}{2}\sigma_2^2 \end{aligned} \tag{4}$$

Because of this important applications in fluid dynamics, equations of the format of (4) have been intensely studied in the literature. We review here the following popular discretization schemes whose performance within a ML estimation problem will be explored by Monte Carlo simulations subsequently:

1. The Peaceman-Rachford Scheme (Strikwerda, 2004, c. 7)

Let us start with a definition of the grid with space and time coordinates $(x_{1,j}, x_{2,l}, t_i)$. Grid points are defined by $x_{1,j} = x_{1,0} + jh_1,\ j = 0, 1 \ldots N_{x_1};\ x_{2,l} = x_{2,0} + lh_2,\ l = 0, 1, \ldots, N_{x_2}$, and $t_i = i\kappa,\ i = 0, 1, \ldots, N_t$. Finally $u^i_{j,l}$ denotes the approximation of the density f at grid points $(x_{1,0} + jh_1,\ x_{2,0} + lh_2,\ i\kappa)$. Here h_1, h_2 and κ are the constant distances between grid points along the two space axes and the time axis. The Peaceman-Rachford scheme evaluates the second-order derivatives in one space direction implicitly (i.e. forward in time) and the second one explicitly (backward in time) within a sequence of time steps of length $\kappa/2$. The cross-derivative is always evaluated explicitly. We denote the finite difference approximations of the second derivatives by:

$$\delta^2_{x_1} = \frac{1}{h^2}(u^i_{j+1,l} - 2u^i_{j,l} + u^i_{j-1,l}) \tag{5}$$

for given coordinates i and l and analogously for $\delta^2_{x_2}$, while

$$\delta_{x_1}\delta_{x_2} = \frac{u^i_{j+1,l+1} - u^i_{j+1,l-1} - u^i_{j-1,l+1} + u^i_{j-1,l-1}}{4h_1h_2} \tag{6}$$

is the finite difference approximation of the cross-derivative. The resulting approximation of eq.(4) over a time interval of length κ requires the following operations:

$$\begin{aligned} \frac{u^{i+\frac{1}{2}}_{j,l} - u^i_{j,l}}{\kappa/2} &= b_{11}\delta^2_{x_1}u^{i+\frac{1}{2}}_{j,l} + 2b_{12}\delta_{x_1}\delta_{x_2}u^i_{j,l} + b_{22}\delta^2_{x_2}u^i_{j,l}, \\ \frac{u^{i+1}_{j,l} - u^{i+\frac{1}{2}}_{j,l}}{\kappa/2} &= b_{11}\delta^2_{x_1}u^{i+\frac{1}{2}}_{j,l} + 2b_{12}\delta_{x_1}\delta_{x_2}u^{i+\frac{1}{2}}_{j,l} + b_{22}\delta^2_{x_2}u^{i+1}_{j,l}. \end{aligned} \tag{7}$$

Rearranging, we arrive at two tridiagonal systems of equations that can easily be solved in an iterative way:

$$\begin{aligned} (1 - \frac{\kappa}{2}b_{11}\delta^2_{x_1})u^{i+\frac{1}{2}}_{j,l} &= (1 + \frac{\kappa}{2}b_{22}\delta^2_{x_2})u^i_{j,l} + \kappa b_{12}\delta_{x_1}\delta_{x_2}u^i_{j,l} \\ (1 - \frac{\kappa}{2}b_{22}\delta^2_{x_2})u^{i+1}_{j,l} &= (1 + \frac{\kappa}{2}b_{11}\delta^2_{x_1})u^{i+\frac{1}{2}}_{j,l} + \kappa b_{12}\delta_{x_1}\delta_{x_2}u^{i+\frac{1}{2}}_{j,l}. \end{aligned} \tag{8}$$

This scheme is known to be of first-order accuracy in general, while it would be second-order accurate in the absence of the mixed term, i.e. with $b_{12} = 0$.

2. The Splitting Scheme of Yanenko *et al.* (1971)

   Pretty much at the same time when Western authors developed various ADI schemes, Russian mathematicians pursued closely related avenues under the heading of splitting schemes or methods of fractional steps. An approach proposed for the bivariate case case with mixed derivatives resembles closely the Peaceman-Rachford scheme, but in its two half-steps takes care of only one space direction, respectively. Computationally, this is only slightly different from the previous scheme.

   It amounts to the sequence:

$$\begin{aligned} \frac{u_{j,l}^{i+\frac{1}{2}} - u_{j,l}^{i}}{\kappa} &= b_{11}\delta_{x_1}^2 u_{j,l}^{i+\frac{1}{2}} + b_{12}\delta_{x_1}\delta_{x_2} u_{j,l}^{i} \\ \frac{u_{j,l}^{i+1} - u_{j,l}^{i+\frac{1}{2}}}{\kappa} &= b_{22}\delta_{x_2}^2 u_{j,l}^{i+1} + b_{12}\delta_{x_1}\delta_{x_2} u_{j,l}^{i+\frac{1}{2}} \end{aligned} \tag{9}$$

   The monograph by Yanenko *et al.* (1971) notes that stability and convergence of this scheme can be proved, but does not provide results on the order of accuracy of this approximation.

3. The McKee-Mitchell Scheme

   This is an alternative popular ADI scheme that goes back to McKee and Mitchell (1970) and has been further elaborated on recently by Craig and Snyd (1988) and McKee *et al.* (1996). The pertinent sequence of operations is:

$$\begin{aligned} (1 - \lambda\kappa b_{11}\delta_{x_1}^2) u_{j,l}^{i+1^*} &= (1 + \kappa(1-\lambda) b_{11}\delta_{x_1}^2 + \kappa b_{22}\delta_{x_2}^2 \\ &\quad + \frac{1}{2}\kappa b_{12}\delta_{x_1}\delta_{x_2}) u_{j,l}^{i} \\ (1 - \lambda\kappa b_{22}\delta_{x_2}^2) u_{j,l}^{i+1} &= u_{j,l}^{i+1^*} - \lambda\kappa b_{22}\delta_{x_2}^2 u_{j,l}^{i} \end{aligned} \tag{10}$$

   In this scheme $i + 1^*$ is an intermediate step and λ is a parameter that can be adjusted to provide for unconditional stability.

   In the present case of two space dimensions, $\lambda \geq 0.5$ guarantees unconditional stability. In the absence of mixed derivatives, the choice of $\lambda = 0.5$ yields a scheme of second-order accuracy both in time and space whereas both with a higher choice of λ or in the presence of mixed terms its accuracy

is of lower order. In our Monte Carlo simulations below, we report results for the choice of $\lambda = 0.5$. We have experimented with higher values but did only find a tendency of deterioration of parameter estimates.

All three schemes will be used within a numerical ML estimation exercise in the next section. Because of certain numerical problems for very high absolute values of ρ outlined below, we also tried some refinements of the baseline ADI and fractional steps methods. One for which we report results is a correction of the cross derivative term proposed recently by Bouchut and Frid (2006). The potential problem with the cross derivatives is that they give rise to the negative entries $-u^i_{j+1,l-1}$ and $-u^i_{j-1,l+1}$ that constitute a thread to the positivity of our numerical solutions. As one can see, this thread materializes itself for high degrees of (positive or negative) correlation, cf. Figs. 1 and 2. Bouchut and Frid propose to add a correction of higher order[1]:

$$\begin{aligned}\delta^+_{x_1}\delta^-_{x_1}\delta^+_{x_2}\delta^-_{x_2}\tilde{b}_{12} = (&u_{j+1,l+1} - 2u_{j,l+1} + u_{j-1,l+1}\\ &- 2u_{j+1,l} + 4u_{j,l} - 2u_{j-1,l}\\ &+ u_{j+1,l-1} - 2u_{j,l-1} + u_{j-1,l-1})\tilde{b}_{12},\end{aligned} \tag{11}$$

that should neutralize these entries. Here, $\tilde{b}_{12}$ denotes an appropriate transformation of the original cross-derivative term for which they found the optimal choice to be $\tilde{b}_{12} = |b_{12}|$. In our Monte Carlo simulations reported in the next section, we adopt this approach alongside with the three baseline methods introduced above, i.e. we consider the performance of each of these schemes with and without the Bouchut and Frid correction. Fig. 1 and 2 illustrate the influence of the mixed derivatives. In Fig. 1, the set of parameters is $\sigma_1 = 0.2$, $\sigma_2 = 0.1$ and $\rho = 0.5$. With this moderate level of dependency, positivity of solutions seems not endangered and, indeed, we can hardly distinguish between the numerical solutions produced by different schemes. Fig. 2 has the same set of parameters except for a much higher correlation of $\rho = 0.95$. Here the potential problem becomes apparent: All methods generate a dent of the distribution where it should flatten out because of the influence of large negative entries from the mixed term. Again, different schemes lead to virtually identical densities. Since the correction scheme by Bouchut and Frid explicitly aims at removing the negative dent, it seems worthwhile to explore its effect on estimation results. However, its performance

[1] $\delta^+_{x_1}$ and $\delta^-_{x_1}$ stand for the finite differences $\delta^+_{x_1} = u_{j+1,l} - u_{j,l}$, $\delta^-_{x_1} = u_{j,l} - u_{j-1,l}$ and analogously for $\delta^+_{x_2}$, $\delta^-_{x_2}$.

turns out to be disappointing as we will see below.[2] We have also tried another proposed remedy, smoothing the solution by time averaging as proposed in Craig and Snyd (1988) for the scheme of McKee *et al.* and in Strikwerda (2004) for the Peaceman-Rachford scheme. As it turned out, this refinement had virtually no effect on parameter estimates so that we abstain from giving details (numbers would be almost exactly identical to those obtained for the respective original schemes in Table 1).

Figs. 1 and 2 about here

4 Monte Carlo Results for Bivariate SDEs

4.1 Bivariate Diffusion

Now we turn to the results of this first set of Monte Carlo experiments using the three finite difference schemes presented in the previous Section. Table 1 reports sample statistics of parameter estimates for the case $\sigma_1 = 0.2, \sigma_2 = 0.1$ and ρ from the set $\{-0, .95, -0.9, -0.5, 0, 0.5, 0.9, 0.95\}$. We have dealt with the negative dent of the numerical density in a hands-on way by replacing negative entries for conditional probabilities by a very small positive number. The hope is that these cases might be rare (as the true density would assume very small positive numbers around the dent anyway) and, therefore, should not affect dramatically our estimation results. Note that for our bivariate Brownian motion, we can actually provide a closed-form solution so that we can also estimate the parameters by exact maximum likelihood (EML). We perform simulations for time series of length $T = 200$ and $T = 500$ to see whether the expected improvement with sample size can be observed (which turns out to be the case). Here and in the following tables, we report the mean parameter estimate across the Monte Carlo repetitions together with its finite sample standard error (FSSE) and the root-mean squared error (RMSE) as evaluated against the "true" parameter.

We start with some very good news: The numerical approach provides estimates whose statistical properties are almost completely equivalent to those from EML for moderate levels of correlation (i.e. $\rho = -0.5, 0$ and 0.5 in our simulation). In these cases, all methods give virtually the same accuracy in parameter estimation

[2] As pointed out by F. Bouchut, the cross derivative correction term might have different implications for stability when applied together with different finite difference schemes. Unfortunately, we cannot use it in our context together with the particular scheme applied by Bouchut and Frid (2006) as the latter is not unconditionally stable.

and except for a slight bias for the estimates of ρ are practically identical to those obtained with EML. For higher positive or negative correlation the estimates of σ_1 and σ_2 have still the same quality for EML like numerical ML while the standard errors for ρ increase somewhat. In these cases we also find small differences in the efficiency of different schemes with the algorithm of McKee *et al.* (MK) dominating those by Peaceman-Rachford (PR) and Yanenko (SP). However, as it turns out, using the correction term by Bouchut and Frid (2006) strongly deteriorates the results for ρ, for true values $|\rho| \geq 0.9$. The correction is, therefore, achieved with a strong bias in the mixed term while our hands-on approach of eliminating any negative entries gives much more satisfactory results with a very moderate bias. The reason for this poor performance might be that the correction is not tailor-made for the present discretisation schemes. Since the scheme developed by McKee *et al.* (1996) has a slight, but consistent advantage over the other methods, we will concentrate on the later in our subsequent exercises and applications.

Table 1 about here

4.2 A Nonlinear Simultaneous System

In the second set of Monte Carlo simulations, we allow for drift functions of highly nonlinear form:

$$\begin{aligned} \mu_1(x_1) &= a_1 + a_2 x_1 + a_3 x_1^2 + a_4 x_1^3 + c_1 y_1, \\ \mu_2(x_2) &= b_1 + b_2 x_2 + b_3 x_2^2 + b_4 x_2^3 + d_1 x_1. \end{aligned} \tag{12}$$

In economics, the study of such nonlinear drift functions has been motivated by formalizations of catastrophe theory (Rheinlaender and Steinkamp, 2004; Barunik and Vosvrda, 2009). Creedy *et al.* (1996) have estimated the parameters of a similarly flexible model for univariate exchange rate data using the ergodic distribution of the so-defined diffusion process. This approach, however, disregards the transitional dynamics and assumes that the data at hand represents a statistical equilibrium which would only hold if the sample size were sufficiently large. Our approach, in contrast, allows to take stock of the transitional effects. Motivated by the related approach of Lux (2012), we investigate one particular scenario: a bivariate process in which one of the variables is driven by a bimodal drift function (also known as double-well model) whereas the second one obeys a unimodal Ornstein-Uhlenbeck dynamics. In addition, we allow for cross-dependencies in the drift as well as for different levels of correlation in the innovations (i.e. the

diffusion terms). The Fokker-Planck equation for the time development of the probability density is now given by:

$$\begin{aligned}\frac{\partial f}{\partial t} = &-\frac{\partial}{\partial x_1}\left[\mu_1(X)f\right] - \frac{\partial}{\partial x_2}\left[\mu_2(X)f\right] \\ &+ \left(b_{11}\frac{\partial^2}{\partial x_1^2} + b_{12}\frac{\partial^2}{\partial x_1 \partial x_2} + b_{22}\frac{\partial^2}{\partial x_2^2}\right) f,\end{aligned} \tag{13}$$

where $X = (x_1, x_2)'$ and b_{11}, b_{12} and b_{22} are defined as in eq.(4). Because of the slight advantage of the scheme proposed by Mckee *et al.* (MK henceforth), and also because of computational concerns, we confine our interest in the following to this particular algorithm. Note that computational cost is higher in the case of system (12) as the Fokker-Planck equation has to be solved numerically for every pair of adjacent observations (as the drift depends on the current realizations of x_1 and x_2) while it had only to be integrated once in the pure diffusion model for each set of parameters. The necessary generalization of the scheme by McKee *et al.* is obtained by adding appropriate convective (i.e. drift-related) terms in eq.(10). This amounts to replacing $\kappa b_{11}\delta_{x_1}^2$ in eq.(10) by $\kappa\left(b_{11}\delta_{x_1}^2 - \mu_1(x_{1,j}, x_{2,l})\delta_{x_1}\right)$ with $\delta_{x_1} = \frac{1}{h_1}(u^i_{j+0.5h_1,l} - u^i_{j-0.5h_1,l})$ and analogously for δ_{x_2}. McKee *et al.* (1996) note that this scheme is of first order accuracy for the general case of a parabolic system with mixed derivative terms and convective (drift) components. Note, however, that their proof assumes linear drift functions so that our case is still more complicated and their results do not apply immediately. Given the absence of strong theoretical results, it is worthwhile to remark that no problems of convergence have been encountered in our simulations.

To reduce the number of parameters, we set $a_1 = a_3 = b_1 = b_3 = b_4 = 0$ which leaves the parameter set $\{a_2, a_4, c_1, b_2, d_1, \sigma_1, \sigma_2, \rho\}$ to be estimated. We keep $a_2 = 1.5$, $a_4 = -3.0$ and $b_2 = -1.5$ constant as well as $\sigma_1 = 1$, $\sigma_2 = 0.5$ over all simulations for the baseline scenario of an intrinsic double-well dynamics of variable x_1 and a mean-reverting unimodal process for x_2. The correlation ρ is varied across the set $\{-0.9, -0.5, 0, 0.5, 0.9\}$ and combined with different cross-dependencies in the drift functions. In scenario 1, cross-dependencies are absent ($c_1 = d_1 = 0$), in scenario 2, a uni-directional influence is assumed ($c_1 = 10, d_1 = 0$), and in scenario 3 both variables influence each other ($c_1 = -10, d_1 = 10$). Table 2 exhibits the results of Monte Carlo simulations using 200 random repetitions for each scenario with $T = 200$ and $T = 500$ observations. While finite sample standard errors and root mean squared errors are higher for the case of seven parameters to be estimated than for the pure diffusion, results

indicate that the means over all samples appear to converge to the true parameter values, with the only exception of a_2 in the third scenario with $\rho = 0.9$. In particular, despite the relatively small sample size, the bimodal and unimodal character of the two drift functions are typically recovered by the parameter estimates, and the estimation procedure is able to distinguish between correlation in innovations and cross-dependencies in the drift term with satisfactory reliability. Comparison between the two sample sizes ($T = 200$ and $T = 500$) shows mostly nice agreement with $\sqrt{T}$ consistency.

Table 2 about here

5 Numerical Maximum Likelihood for Trivariate SDEs

Although the computational demands of our approach increase quite significantly with the dimension of the time series, parallelization of the computational task still allows us to handle trivariate systems. To get some evidence on the performance of the numerical ML approach, we again conduct simulations for a pure diffusion process which now is characterized by six parameters $\{\sigma_1, \sigma_2, \sigma_3, \rho_{12}, \rho_{23}, \rho_{13}\}$. One might note that the computational demands are much lower for the pure diffusion than for models with non-zero drift components: since the development in time of the transient pdf is independent of the state variables $\{x_{1,t}, x_{2,t}, x_{3,t}\}$, we only need to solve the Fokker-Planck equation numerically once for each set of parameters. Hence, the necessary iterations of the numerical scheme only depend on the number of iterations we need for the convergence of the optimization routine for the likelihood function (we use a combination of the Newton-Raphson and Broyden-Fletcher-Goldfarb-Shanno algorithms). In contrast, with state-dependent drift functions, the number of iterations would also depend on the number of observations (as each set of initial conditions would have to be taken into account through a new numerical solution of the Fokker-Planck equation). This feature allows us to perform a similar Monte Carlo study as with the bivariate diffusion: We consider sample sizes of $T = 200$ and $T = 500$ observations and report results from 200 random repetitions for each scenario. Because of the very promising previous experiences with the MK scheme, we stick to this finite difference algorithm and also dispense with any additional correction for mixed derivatives. The MK scheme in three space dimensions is given by the

following sequence of operations:

$$
\begin{aligned}
(1-\lambda\kappa b_{11}\delta_{x_1}^2)u_{j,l,m}^{i+1^*} &= (1-\lambda\kappa b_{11}\delta_{x_1}^2 + \sum_{r=1}^{3}\kappa b_{rr}\delta_{x_r}^2 \\
&\quad + \frac{1}{2}\sum_{r=2}^{3}\sum_{s=1}^{2}\kappa b_{rs}\delta_{x_r}\delta_{x_s})u_{j,l,m}^{i} \qquad (14)\\
(1-\lambda\kappa b_{22}\delta_{x_2}^2)u_{j,l,m}^{i+1^{**}} &= u_{j,l,m}^{i+1^*} - \lambda\kappa b_{22}\delta_{x_2}^2 u_{j,l,m}^{i} \\
(1-\lambda\kappa b_{33}\delta_{x_3}^2)u_{j,l,m}^{i+1} &= u_{j,l,m}^{i+1^{**}} - \lambda\kappa b_{33}\delta_{x_3}^2 u_{j,l,m}^{i},
\end{aligned}
$$

where $u_{j,l,m}^{i}$ is the trivariate density evaluated at $x_{1,j}$, $x_{2,l}$, $x_{3,m}$. Allowing for convective terms (as in our empirical application below) could be accommodated by adding the finite difference approximations of the drift terms as described in section 4.2. Setting $\lambda \geq 2/3$ should guarantee unconditional stability for the case of a pure diffusion as well as with linear convective terms added.

Table 3 exhibits the results for a selection of sets of parameter values. While we have kept the variances unchanged, the correlations cover a range of possibilities (note that positivity of the covariance matrix restricts the number of possible combinations). Results are overall supportive, but also show some surprising features. Again, we can compare our numerical results with those from exact ML as the transient density of this simple process is, of course, given by a trivariate Normal distribution. The first surprising feature of Table 3 is that we see a hierarchy of parameters: while σ_1, σ_2 and ρ_{12} tend to be estimated as accurately under the numerical scheme as with EML, results for σ_3, ρ_{23} and ρ_{13} show somewhat larger variations. In particular, the later show a small bias that does not seem to vanish with increasing T, and higher FFSE and RMSE than the former. Furthermore, while the first set of parameters appears to be estimated with $\sqrt{T}$ consistency, the errors of the second set of parameters appear to decrease more slowly. Closer inspection reveals some of the potential reasons for this behavior. In particular, while the first three parameters have MC distributions that are well-approximated by a Normal distribution, the others have more asymmetric distributions with a long tail on one side. For instance, for the first set of parameters, σ_3 is right-skewed while ρ_{23} and ρ_{13} have a heavy left tail. Furthermore, the deviation of σ_3, on the one hand, and ρ_{23} and ρ_{13}, on the other hand, are negatively correlated. This speaks for some sort of near-collinearity of these variables. Our conjecture is that the elimination of the mixed term from the auxiliary steps of the MK scheme might be responsible for this distortion. It might, therefore, be worthwhile to also experiment with alternative schemes that also take into account the influence of

the mixed term beyond the first fractional step. It is, however, also worthwhile to mention that the correlations are estimated with about the same precision under the numerical ML scheme as with EML. The only parameter where a certain inferiority of the numerical approach shows up consistently, is σ_3.

Table 3 about here

6 An Empirical Application

Since despite certain problems, the numerical schemes get typically very close to EML estimation results, we feel encouraged to proceed with an empirical application. For an illustration of our methodology we revisit the data investigated by Lux (2012). In this paper, various versions of a continuous-time asset pricing model are estimated using weekly short-and medium run sentiment data obtained from *animusX-Investor sentiment* (http://animusX.de) together with German DAX data at the same frequency. Lux (2012) has modeled the dynamics of both sentiment indices by a opinion formation process leading to a highly nonlinear bivariate diffusion. The price dynamics were assumed to follow a Wiener process with a drift term depending on sentiment. One drawback of the opinion dynamics studied in this paper was that it could not easily accommodate mixed derivatives in the diffusion term (as the later was already highly nonlinear itself). It turned out that the estimation resulted in some ambiguity concerning the significance of short-run and medium-run sentiment in the drift of the price equation. It was conjectured that such an ambiguity might have been due to the negligence of mixed derivatives which could have been absorbed by the drift function. Our more general setting confirms that this could indeed have been the case. Denoting by x_t short-run sentiment, y_t medium-run sentiment and p_t weekly DAX notations, our estimated trivariate model is:

$$\begin{aligned} dx_t &= \left(a_1 + a_2 x_t + a_3 x_t^2 + a_4 x_t^3 + c_1 y_t + c_2 \Delta p_t\right) dt + \sigma_x dW_1, \\ dy_t &= \left(b_1 + b_2 y_t + b_3 y_t^2 + b_4 y_t^3 + d_1 x_t + d_2 \Delta p_t\right) dt + \sigma_y dW_2, \\ dp_t &= e_0 + e_1 x_t + e_2 y_t + \sigma_p dW_3, \end{aligned} \tag{15}$$

with W_1, W_2, W_3 scalar Wiener processes with correlations ρ_{xy}, ρ_{yp}, and ρ_{xp}, and Δp_t denoting the price change between unit time steps: $\Delta p_t = p_t - p_{t-1}$. Tables 4 to 7 exhibit parameter estimates first for restricted bivariate models for each set

of two of our variables (with all reference to the third deleted) as well as finally of the complete trivariate model. The data used for estimation are the first 150 observations of the animusX data starting in the $29th$ calender week of 2004 plus the pertinent DAX entries.

Tables 4 to 6 show the full bivariate models with estimates of all relevant parameters as well as reduced models obtained after a model reduction exercise. For the later, various sub-models have been estimated and the one with the lowest Akaike information criterion (AIC) has been retained to identify the set of significant parameters. To highlight the relevance of the mixed derivative, the reduced model has been re-estimated also without a covariance term. As we can see, the various models unambiguously support bimodality of x_t (short-run sentiment) which might be explained by speculative exuberance, and unimodality of y_t (medium-run sentiment). We find that there is a consistent influence from y_t on x_t but not vice versa (in harmony with Lux, 2011, 2012). Covariances appear indispensable in all cases. Most importantly, only y_t seemed to have an influence on prices while x_t dropped out consistently for the price equation in the model reduction exercise.

These features are fully confirmed in the estimation of the trivariate model (cf. Table 7). For brevity, we only report those parameters that remained significant in a more parsimonious model under the AIC criterion (model I) as well as restricted models in which each one of the covariances has been set to zero (models II through IV). While Lux (2012) found a crucial inconsistency between his bivariate and trivariate models concerning the determinants of price drift, the present results are mutually consistent: while x_t has the more interesting dynamics, it seems rather autistic and does not drive prices but rather is driven itself by y_t. In contrast, y_t exerts a significant influence on subsequent returns both in the bivariate (Table 4) and the trivariate (Table 6) estimation exercises. This is also in harmony with the findings for a purely statistical discrete-time vector autoregressive model (Lux, 2011).

Similarly as Lux (2012), we use the estimated model to forecast prices out-of-sample. The out-of-sample period covers 192 observations from mid 2007 to January 2011. Table 8 shows root mean-squared errors (RMSEs) of single week and cumulative returns for horizons from one to eight weeks. As can be seen, our forecasts provide a significant improvement against the random walk benchmark at longer horizons, while the model investigated in Lux (2012) was unable to beat a pure random walk at any horizon. Apparently, the consideration of mixed derivatives has improved performance against the restricted (and somewhat dif-

ferent) framework explored before. While the mixed terms would, of course, not provide forecasting power per se, their inclusion in the model may have prevented us from confounding correlation in innovations with causal dependencies in the drift term. Note that restricting covariances to zero is often reflected in larger changes of cross-dependencies in the drift term than in other parameters. This highlights a certain danger of erroneously attributing correlation of innovations to causal dependency if the former is not explicitely included in the model.

Tables 4 through 8 about here

7 Conclusion

Expanding a previous line of research initiated by Poulsen (1999) and Hurn et al. (2009, 2010) we have explored the potential of numerical approximations of the Fokker-Planck equation for maximum likelihood estimation of discretely sampled diffusions. We have pointed out certain requirements like positivity and stability that should be satisfied by numerical schemes to be applied in this context. Monte Carlo simulations of different alternating direction finite difference schemes for bivariate and trivariate series showed very promising results often undistinguishable from the performance of exact maximum likelihood on the base of a closed form solution of the transient density. Even known deficiencies like negative dents for high values of the mixed derivatives proved to cause no harm to the estimates under a hands-on elimination of problematic entries. Our empirical application suggests that a full-fletched multivariate system of differential equations might capture both causal relationships between variables as well as correlation in innovations and could, thereby, reveal a rich spectrum of co-evolutionary dynamics of the quantities of interest. Since we have only been experimenting with a limited range of possible finite difference schemes, there is ample scope for further work. In particular, our mixed results for the precision of estimates of different parameters in the trivariate case might suggest to use a more symmetric scheme than the one proposed by McKee *et al.* (1996) that has a somewhat preferential treatment of the first space dimension. Given potential instability of a range of alternative schemes (which has not been encountered in all our examples), the effect of upwinding for numerical stabilization on parameter estimates might also be investigated. Indeed, where results on the precision of a certain scheme are not available, our parameter estimation exercise might shed light on its performance and possible distortions it could give rise to.

References

[1] Ait-Sahalia, Y. (1999). Transition densities for interest rate and other nonlinear diffusions, *Journal of Finance* 54, 1361 – 1395.

[2] Ames, W. (1992). *Numerical Methods for Partial Differential Equations*, 3rd ed. Boston, Academic Press.

[3] Barunik, J. and M. Vosvrda (2009). Can a stochastic cusp catastrophe model explain stock market crashes?, *Journal of Economic Dynamics and Control* 33, 1824 – 1836.

[4] Bouchut, F. and H. Frid (2006). Finite difference schemes with cross derivatives correctors for multidimensional parabolic systems, *Journal of Hyperbolic Differential Equations* 3, 27 – 52.

[5] Craig, I. and A. Snyd (1988). An alternating-direction implicit scheme for parabolic equations with mixed derivatives, *Computational Mathematics and Applications* 16, 341 – 350.

[6] Creedy, J. , J. Lye, and V. Martin (1996). A non-linear model of the real US/UK exchange rate, *Journal of Applied Econometrics* 11, 669 – 686.

[7] Hurn, A., J. Jeisman, and K. Lindsay (2009). Seeing the wood for the trees: a critical evaluation of methods to estimate the parameters of stochastic differential equations, *Journal of Financial Econometrics*, 5, 390 – 455.

[8] Hurn, A., J. Jeisman, and K. Lindsay (2010). Teaching an old dog new tricks: improved estimation of the parameters of stochastic differential equations by numerical solution of the Fokker-Planck equation, Gregorion, G. N. and Pascalau, R., eds., *Financial Econometrics Handbook*, London, Palgrave.

[9] Iacus, S. M. (2008). *Simulation and Inference for Stochastic Differential Equations*, Springer Series in Statistics.

[10] Jensen, B., R. Poulsen (2002). Transition densities of diffusion processes: numerical comparison of approximation techniques, *Journal of Derivatives* 9, 18 – 34.

[11] Kloeden, P. and E. Platen (1992). *Numerical Solutions of Stochastic Differential Equations*, Berlin, Springer.

[12] Lux, T., (2009). Rational forecasts or social opinion dynamics? Identification of interaction effects in a business climate survey, *Journal of Economic Behavior and Organization* 72, 638 – 655.

[13] Lux, T., (2011). Sentiment dynamics and stock returns: the case of the German stock market, *Empirical Economics*, 41, 663 – 679.

[14] Lux, T., (2012). Estimation of an agent-based model of investor sentiment formation of financial markets, *Journal of Economic Dynamic of Control*, 36, 1284 – 1302.

[15] Marchuk, G. (1990). Splitting and alternating direction methods, Ciarlet, P. G. and Lions, J. L., eds. *Handbook of Numerical Analysis*, 1, Amsterdam, North-Holland, 197 – 462.

[16] McKee, S. and A. Mitchell (1970). Alternating direction methods for parabolic equations in two space dimensions with a mixed derivative, *Computer Journal*, 13, 81 – 86.

[17] McKee, S., D. Wall and S. Wilson (1996). An alternating-direction implicit scheme for parabolic equations with mixed derivative and corrective terms, *Journal of Computational Physics* 126, 64 – 76.

[18] Morton, K. and D. Mayers (1994). *Numerical Solution of Partial Differential Equations*, Cambridge University Press.

[19] Poulsen, R., (1999). *Approximate Maximum Likelihood Estimation of Discretely Observed Diffusion Processes*, Manuscript, University of Aarhus.

[20] Rheinlaender, T. and M. Steinkamp (2004). A Stochastic version of Zeeman's market model, *Studies in Nonlinear Dynamics and Econometrics*, 8.

[21] Strikwerda, J. (2004). *Finite Difference Schemes and Partial Differential Equations*, 2nd ed., Philadelphia: Society for Industrial and Applied Mathematics.

[22] Yanenko, N. (1971), *The Method of Fractional Steps: The Solution of Problems of Mathematical Physics in Several Variables*, Berlin, Springer.

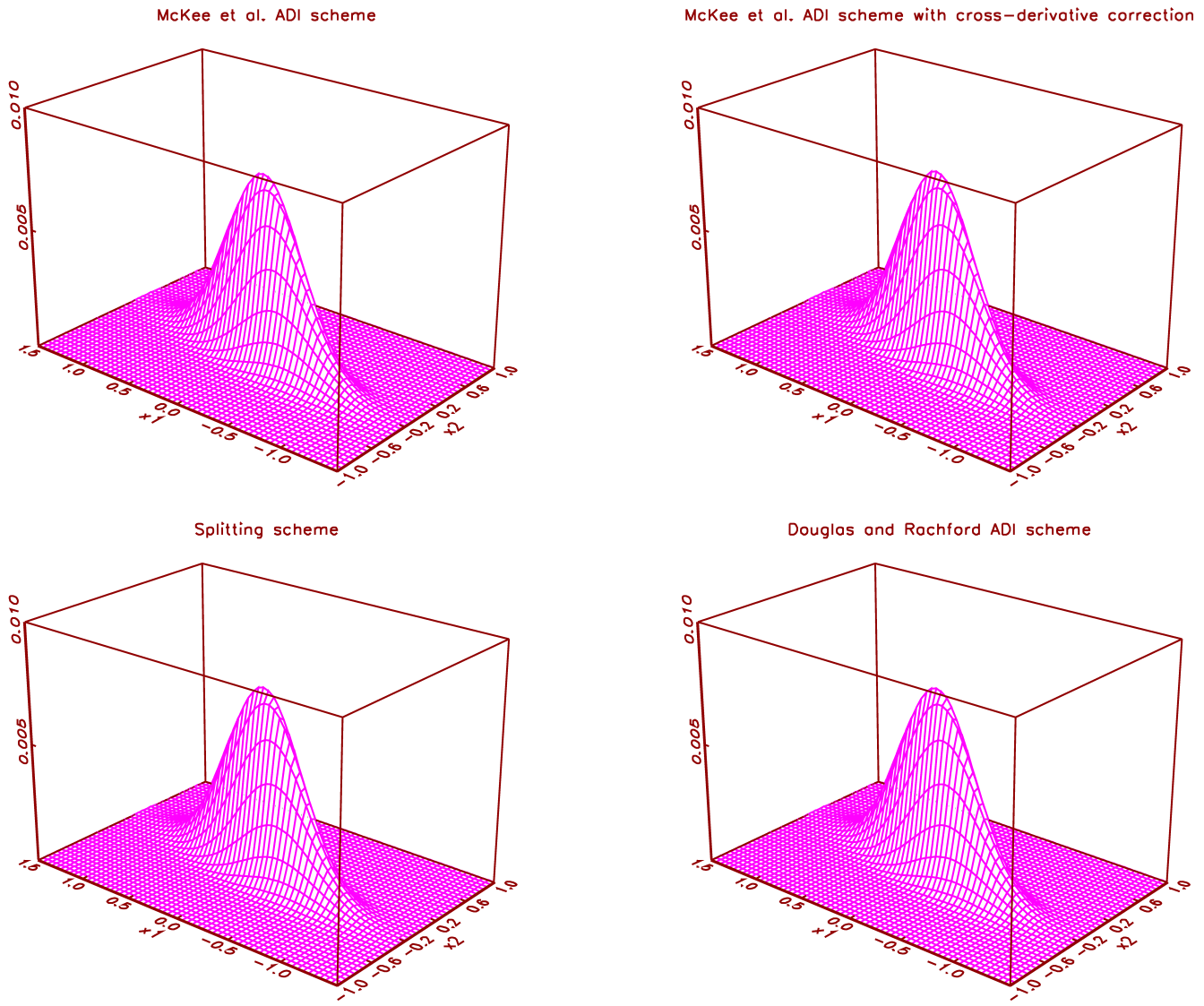

Figure 1: Comparison of numerical approximations of the conditional density of a bivariate diffusion with parameters $\sigma_1 = 0.2$, $\sigma_2 = 0.1$, and $\rho = 0.5$.

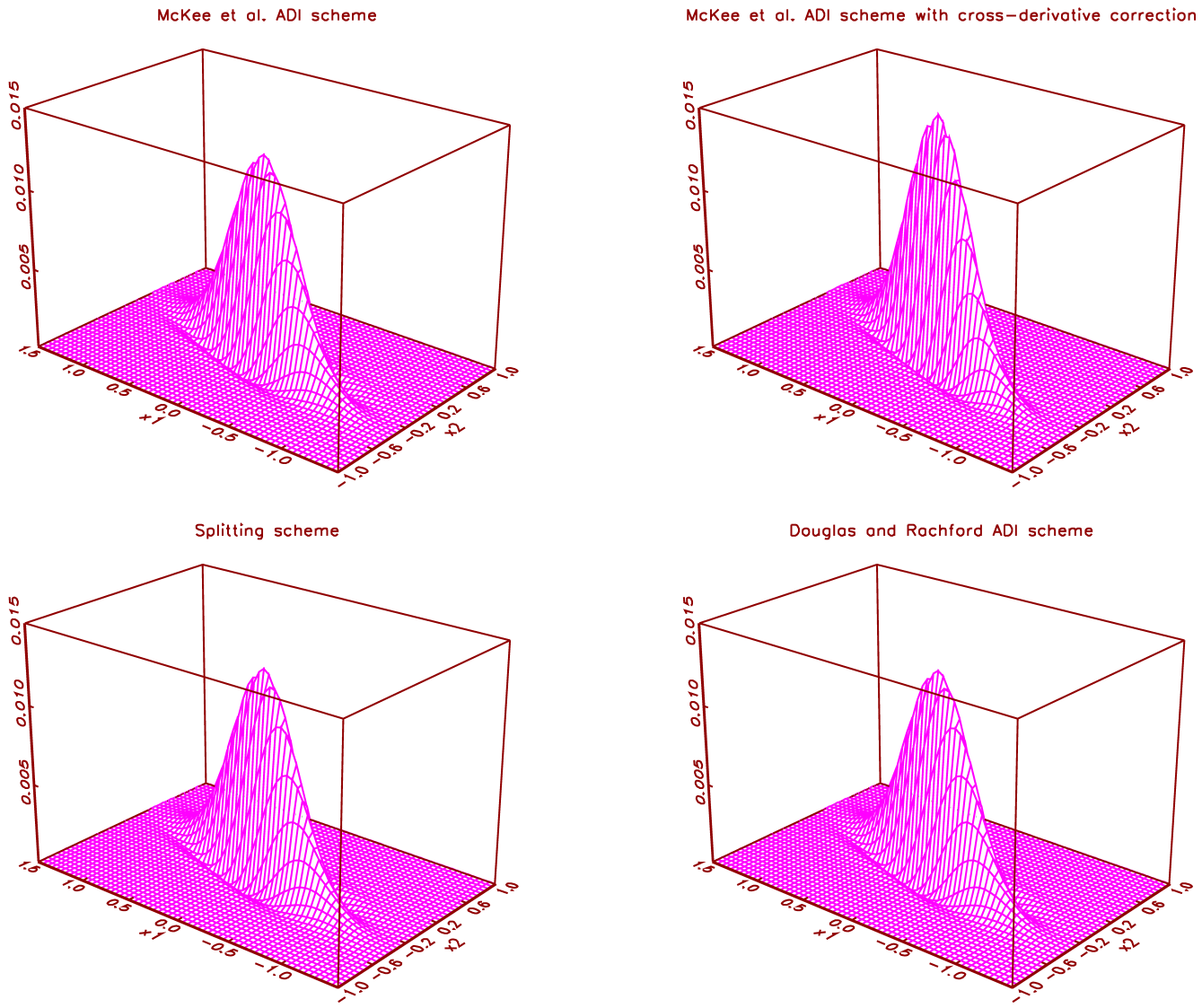

Figure 2: Comparison of numerical approximations of the conditional density of a bivariate diffusion with parameters $\sigma_1 = 0.2$, $\sigma_2 = 0.1$, and $\rho = 0.95$.

Table 1: Monte Carlo Results for Bivariate Diffusion

	$T = 200$							$T = 500$						
	EML	DR	SP	MK	DR_c	SP_c	MK_c	EML	DR	SP	MK	DR_c	SP_c	MK_c
$\rho = -0.950$														
Parameter σ_1														
$mean$	0.201	0.182	0.209	0.205	0.196	0.197	0.186	0.200	0.180	0.207	0.205	0.196	0.196	0.185
$FSSE$	0.010	0.014	0.013	0.012	0.010	0.010	0.009	0.006	0.009	0.009	0.007	0.006	0.006	0.006
$RMSE$	0.010	0.023	0.016	0.013	0.010	0.011	0.017	0.006	0.022	0.011	0.009	0.007	0.007	0.016
Parameter σ_2														
$mean$	0.100	0.092	0.104	0.103	0.098	0.098	0.092	0.100	0.091	0.103	0.102	0.098	0.098	0.092
$FSSE$	0.005	0.013	0.006	0.006	0.005	0.005	0.004	0.003	0.005	0.004	0.004	0.003	0.003	0.003
$RMSE$	0.005	0.015	0.007	0.007	0.005	0.005	0.009	0.003	0.010	0.006	0.005	0.004	0.004	0.008
Parameter $\rho = -0.950$														
$mean$	-0.950	-0.962	-0.948	-0.961	-0.930	-0.902	-0.890	-0.950	-0.957	-0.943	-0.958	-0.930	-0.902	-0.890
$FSSE$	0.007	0.026	0.029	0.022	0.008	0.009	0.009	0.004	0.025	0.024	0.017	0.005	0.005	0.006
$RMSE$	0.007	0.028	0.029	0.024	0.021	0.048	0.061	0.004	0.026	0.025	0.019	0.020	0.048	0.060
$\rho = -0.900$														
Parameter σ_1														
$mean$	0.201	0.194	0.203	0.201	0.198	0.198	0.192	0.200	0.191	0.202	0.201	0.198	0.198	0.192
$FSSE$	0.010	0.013	0.011	0.010	0.010	0.010	0.009	0.006	0.009	0.006	0.007	0.006	0.006	0.006
$RMSE$	0.010	0.014	0.011	0.010	0.010	0.010	0.012	0.006	0.013	0.007	0.007	0.006	0.007	0.010
Parameter σ_2														
$mean$	0.100	0.097	0.101	0.100	0.099	0.099	0.096	0.100	0.095	0.101	0.100	0.099	0.099	0.096
$FSSE$	0.005	0.007	0.005	0.005	0.005	0.005	0.004	0.003	0.004	0.003	0.003	0.003	0.003	0.003
$RMSE$	0.005	0.007	0.005	0.005	0.005	0.005	0.006	0.003	0.006	0.003	0.003	0.003	0.003	0.005
Parameter $\rho = -0.900$														
$mean$	-0.900	-0.889	-0.879	-0.895	-0.884	-0.857	-0.851	-0.900	-0.875	-0.874	-0.892	-0.884	-0.857	-0.852
$FSSE$	0.013	0.036	0.017	0.017	0.014	0.014	0.013	0.008	0.027	0.012	0.011	0.008	0.009	0.008
$RMSE$	0.013	0.037	0.027	0.017	0.021	0.045	0.050	0.008	0.037	0.028	0.014	0.018	0.043	0.049

	$T = 200$							$T = 500$						
	EML	DR	SP	MK	DR_c	SP_c	MK_c	EML	DR	SP	MK	DR_c	SP_c	MK_c
$\rho = -0.500$														
Parameter σ_1														
$mean$	0.201	0.200	0.201	0.200	0.200	0.201	0.200	0.200	0.200	0.201	0.200	0.200	0.201	0.200
$FSSE$	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.006	0.006	0.006	0.006	0.006	0.006	0.006
$RMSE$	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.006	0.006	0.006	0.006	0.006	0.006	0.006
Parameter σ_2														
$mean$	0.100	0.100	0.101	0.100	0.100	0.101	0.100	0.100	0.100	0.101	0.100	0.100	0.101	0.100
$FSSE$	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.003	0.003	0.003	0.003	0.003	0.003	0.003
$RMSE$	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.003	0.003	0.003	0.003	0.003	0.003	0.003
Parameter $\rho = -0.500$														
$mean$	-0.497	-0.492	-0.484	-0.489	-0.491	-0.482	-0.486	-0.500	-0.494	-0.486	-0.491	-0.493	-0.484	-0.488
$FSSE$	0.051	0.049	0.049	0.049	0.049	0.049	0.049	0.031	0.033	0.033	0.033	0.033	0.033	0.033
$RMSE$	0.051	0.049	0.051	0.050	0.050	0.052	0.051	0.031	0.033	0.036	0.034	0.034	0.036	0.035
$\rho = 0.000$														
Parameter σ_1														
$mean$	0.201	0.200	0.201	0.200	0.200	0.201	0.200	0.200	0.200	0.201	0.200	0.200	0.200	0.200
$FSSE$	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.006	0.006	0.006	0.006	0.006	0.006	0.006
$RMSE$	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.006	0.006	0.006	0.006	0.006	0.006	0.006
Parameter σ_2														
$mean$	0.100	0.101	0.101	0.101	0.101	0.101	0.101	0.100	0.101	0.101	0.101	0.101	0.101	0.101
$FSSE$	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.003	0.003	0.003	0.003	0.003	0.003	0.003
$RMSE$	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.003	0.003	0.003	0.003	0.003	0.003	0.003
Parameter $\rho = 0.000$														
$mean$	0.004	0.006	0.005	0.005	0.006	0.005	0.005	-0.001	0.002	0.002	0.002	0.002	0.002	0.002
$FSSE$	0.067	0.069	0.068	0.069	0.069	0.068	0.069	0.043	0.046	0.045	0.046	0.046	0.045	0.046
$RMSE$	0.067	0.069	0.068	0.069	0.069	0.068	0.069	0.043	0.046	0.045	0.045	0.046	0.045	0.045

	$T = 200$							$T = 500$						
	EML	DR	SP	MK	DR_c	SP_c	MK_c	EML	DR	SP	MK	DR_c	SP_c	MK_c
$\rho = 0.500$														
Parameter σ_1														
$mean$	0.201	0.200	0.201	0.200	0.200	0.201	0.200	0.200	0.200	0.201	0.200	0.200	0.201	0.200
$FSSE$	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.006	0.006	0.006	0.006	0.006	0.006	0.006
$RMSE$	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.006	0.006	0.006	0.006	0.006	0.006	0.006
Parameter σ_2														
$mean$	0.100	0.101	0.101	0.101	0.101	0.101	0.101	0.100	0.101	0.101	0.101	0.101	0.101	0.101
$FSSE$	0.005	0.004	0.005	0.004	0.004	0.005	0.004	0.003	0.003	0.003	0.003	0.003	0.003	0.003
$RMSE$	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.003	0.003	0.003	0.003	0.003	0.003	0.003
Parameter $\rho = 0.500$														
$mean$	0.503	0.500	0.491	0.496	0.498	0.488	0.493	0.498	0.498	0.489	0.494	0.496	0.487	0.492
$FSSE$	0.052	0.055	0.054	0.054	0.054	0.053	0.054	0.033	0.034	0.034	0.034	0.034	0.034	0.034
$RMSE$	0.052	0.054	0.055	0.054	0.054	0.054	0.054	0.033	0.034	0.036	0.035	0.034	0.036	0.035
$\rho = 0.900$														
Parameter σ_1														
$mean$	0.201	0.194	0.203	0.201	0.198	0.198	0.192	0.200	0.192	0.202	0.201	0.198	0.198	0.192
$FSSE$	0.010	0.014	0.011	0.011	0.010	0.010	0.009	0.006	0.010	0.007	0.007	0.006	0.006	0.006
$RMSE$	0.010	0.015	0.011	0.011	0.010	0.010	0.012	0.006	0.013	0.007	0.007	0.006	0.007	0.010
Parameter σ_2														
$mean$	0.100	0.098	0.102	0.101	0.099	0.099	0.096	0.100	0.096	0.101	0.101	0.099	0.099	0.096
$FSSE$	0.005	0.007	0.005	0.005	0.005	0.005	0.004	0.003	0.005	0.003	0.003	0.003	0.003	0.003
$RMSE$	0.005	0.008	0.005	0.005	0.005	0.005	0.006	0.003	0.006	0.004	0.003	0.003	0.003	0.005
Parameter $\rho = 0.900$														
$mean$	0.900	0.888	0.879	0.895	0.884	0.857	0.851	0.899	0.872	0.873	0.891	0.884	0.858	0.852
$FSSE$	0.014	0.036	0.020	0.018	0.014	0.014	0.013	0.009	0.031	0.012	0.012	0.008	0.009	0.008
$RMSE$	0.013	0.038	0.029	0.018	0.021	0.045	0.050	0.009	0.042	0.030	0.015	0.018	0.043	0.049

	$T = 200$							$T = 500$						
	EML	DR	SP	MK	DR_c	SP_c	MK_c	EML	DR	SP	MK	DR_c	SP_c	MK_c
$\rho = 0.950$														
Parameter σ_1														
mean	0.201	0.182	0.208	0.205	0.196	0.197	0.186	0.200	0.181	0.208	0.205	0.196	0.196	0.186
FSSE	0.010	0.013	0.014	0.011	0.010	0.010	0.009	0.006	0.009	0.009	0.008	0.006	0.006	0.006
RMSE	0.010	0.022	0.016	0.013	0.010	0.011	0.017	0.006	0.021	0.012	0.010	0.007	0.007	0.015
Parameter σ_2														
mean	0.100	0.093	0.104	0.103	0.098	0.098	0.093	0.100	0.091	0.104	0.102	0.098	0.098	0.092
FSSE	0.005	0.007	0.008	0.006	0.005	0.005	0.004	0.003	0.007	0.004	0.004	0.003	0.003	0.003
RMSE	0.005	0.010	0.009	0.006	0.005	0.005	0.008	0.003	0.011	0.006	0.005	0.004	0.004	0.008
Parameter $\rho = 0.950$														
mean	0.950	0.964	0.947	0.960	0.931	0.903	0.890	0.950	0.959	0.944	0.960	0.931	0.903	0.890
FSSE	0.007	0.027	0.033	0.023	0.008	0.009	0.009	0.004	0.024	0.025	0.019	0.005	0.005	0.006
RMSE	0.007	0.030	0.033	0.026	0.021	0.048	0.060	0.004	0.026	0.025	0.021	0.020	0.047	0.060

Note: The table reports the statistics of 200 Monte Carlo runs of each parameter set with sample sizes $T = 200$ and $T = 500$. Results obtained for exact maximum likelihood (EML) are compared with those of six different numerical approximations to the likelihood function: The Douglas-Rachford algorithm (DR), the splitting scheme of Yanenko *et al.* (SP), the ADI method of McKee *et al.* (MK) as well as adaptations of these schemes using the cross derivatives correction proposed by Bouchut and Frid (DR_c, SP_c and MK_c). Chosen grid increments are $\kappa = 1/12$, and the boundaries along the x and y dimensions are chosen as the next higher multiple of 0.1 beyond the maximum of the absolute observations of x_t and y_t, respectively, and equidistant increments h_1 and h_2 are fixed so that $N_x = N_y = 50$. In this application, the origin of the spatial grid is readjusted with the previous observation serving as the origin for the next iteration of the ML procedure.

Table 2: Monte Carlo results for Bivariate Model with Nonlinear Drift Components

	$T = 200$								$T = 500$							
	a_2	a_4	c_1	b_2	d_1	σ_1	σ_2	ρ	a_2	a_2	c_1	b_2	d_1	σ_1	σ_2	ρ
Parameter set 1																
True	1.5	-3.0	0	-1.5	0	1.0	0.5	-0.9	1.5	-3.0	0	-1.5	0	1.0	0.5	-0.9
Mean	1.752	-3.311	0.618	-2.126	-0.116	1.030	0.518	-0.935	1.538	-3.066	0.442	-1.818	-0.042	1.033	0.517	-0.916
FSSE	0.912	0.880	1.968	1.005	0.431	0.062	0.035	0.043	0.687	0.670	1.313	0.639	0.250	0.041	0.022	0.039
RMSE	0.944	0.931	2.058	1.508	0.445	0.068	0.039	0.056	0.686	0.672	1.382	1.037	0.252	0.053	0.028	0.042
True	1.5	-3.0	0	-1.5	0	1.0	0.5	-0.5	1.5	-3.0	0	-1.5	0	1.0	0.5	-0.5
Mean	1.615	-3.235	0.099	-1.769	-0.035	1.017	0.511	-0.474	1.411	-2.978	-0.018	-1.574	-0.012	1.013	0.509	-0.473
FSSE	0.787	0.906	1.090	0.598	0.265	0.055	0.030	0.053	0.553	0.605	0.712	0.316	0.140	0.034	0.019	0.035
RMSE	0.793	0.934	1.092	0.973	0.267	0.058	0.031	0.059	0.559	0.604	0.711	0.655	0.141	0.037	0.021	0.045
True	1.5	-3.0	0	-1.5	0	1.0	0.5	0	1.5	-3.0	0	-1.5	0	1.0	0.5	0
Mean	1.534	-3.207	0.148	-1.734	-0.002	1.019	0.513	0.004	1.373	-2.964	0.085	-1.572	-0.015	1.014	0.511	0.003
FSSE	0.895	1.031	1.081	0.563	0.215	0.056	0.027	0.074	0.593	0.660	0.805	0.290	0.130	0.034	0.018	0.049
RMSE	0.894	1.049	1.088	0.925	0.214	0.059	0.030	0.074	0.605	0.659	0.808	0.641	0.131	0.037	0.021	0.049
True	1.5	-3.0	0	-1.5	0	1.0	0.5	0.5	1.5	-3.0	0	-1.5	0	1.0	0.5	0.5
Mean	1.618	-3.305	0.172	-1.695	0.016	1.018	0.512	0.478	1.426	-3.014	0.115	-1.538	-0.020	1.013	0.510	0.476
FSSE	0.791	0.905	1.078	0.557	0.223	0.055	0.024	0.058	0.539	0.590	0.704	0.334	0.141	0.034	0.017	0.038
RMSE	0.798	0.953	1.089	0.890	0.223	0.058	0.027	0.062	0.543	0.589	0.711	0.633	0.142	0.036	0.020	0.045
True	1.5	-3.0	0	-1.5	0	1.0	0.5	0.9	1.5	-3.0	0	-1.5	0	1.0	0.5	0.9
Mean	1.699	-3.444	-0.179	-1.960	0.056	1.033	0.518	0.931	1.459	-3.052	-0.285	-1.739	0.006	1.031	0.518	0.911
FSSE	0.955	0.946	1.908	0.944	0.365	0.061	0.033	0.042	0.651	0.626	1.237	0.624	0.256	0.040	0.022	0.037
RMSE	0.974	1.043	1.912	1.345	0.369	0.069	0.037	0.053	0.650	0.626	1.267	0.966	0.255	0.051	0.028	0.039

	$T = 200$								$T = 500$							
	a_2	a_4	c_1	b_2	d_1	σ_1	σ_2	ρ	a_2	a_2	c_1	b_2	d_1	σ_1	σ_2	ρ
Parameter set 2																
True	1.5	-3.0	10	-1.5	0	1.0	0.5	-0.9	1.5	-3.0	10	-1.5	0	1.0	0.5	-0.9
Mean	1.601	-3.285	11.253	-1.745	0.015	1.026	0.508	-0.928	1.527	-3.138	10.717	-1.597	0.002	1.022	0.505	-0.914
FSSE	0.890	0.699	1.713	0.739	0.212	0.072	0.039	0.065	0.581	0.470	1.040	0.401	0.111	0.049	0.022	0.045
RMSE	0.894	0.754	2.119	1.048	0.212	0.076	0.039	0.070	0.580	0.489	1.261	0.719	0.111	0.054	0.023	0.047
True	1.5	-3.0	10	-1.5	0	1.0	0.5	-0.5	1.5	-3.0	10	-1.5	0	1.0	0.5	-0.5
Mean	1.574	-3.123	10.527	-1.723	0.015	1.024	0.510	-0.479	1.493	-2.992	10.210	-1.554	-0.003	1.017	0.508	-0.473
FSSE	0.776	0.652	1.493	0.602	0.193	0.062	0.032	0.057	0.500	0.389	0.798	0.372	0.106	0.039	0.020	0.039
RMSE	0.777	0.662	1.580	0.940	0.193	0.066	0.034	0.061	0.499	0.389	0.824	0.667	0.106	0.042	0.022	0.047
True	1.5	-3.0	10	-1.5	0	1.0	0.5	0	1.5	-3.0	10	-1.5	0	1.0	0.5	0
Mean	1.572	-3.154	10.618	-1.749	0.026	1.022	0.513	-0.010	1.497	-3.040	10.350	-1.554	0.002	1.017	0.511	-0.010
FSSE	0.765	0.629	1.670	0.693	0.195	0.060	0.029	0.080	0.452	0.336	0.982	0.418	0.118	0.036	0.019	0.050
RMSE	0.766	0.646	1.776	1.019	0.196	0.063	0.032	0.080	0.451	0.337	1.040	0.693	0.117	0.040	0.022	0.051
True	1.5	-3.0	10	-1.5	0	1.0	0.5	0.5	1.5	-3.0	10	-1.5	0	1.0	0.5	0.5
Mean	1.553	-3.131	10.521	-1.710	0.023	1.018	0.512	0.462	1.493	-3.037	10.248	-1.554	0.001	1.017	0.509	0.457
FSSE	0.728	0.555	1.908	0.752	0.216	0.055	0.026	0.068	0.443	0.339	1.210	0.496	0.139	0.035	0.018	0.043
RMSE	0.728	0.569	1.973	1.032	0.216	0.058	0.028	0.077	0.442	0.341	1.233	0.743	0.138	0.039	0.020	0.061
True	1.5	-3.0	10	-1.5	0	1.0	0.5	0.9	1.5	-3.0	10	-1.5	0	1.0	0.5	0.9
Mean	1.904	-3.486	10.953	-2.162	0.103	1.023	0.517	0.945	1.661	-3.190	10.251	-1.985	0.076	1.030	0.515	0.926
FSSE	0.845	0.624	2.655	1.179	0.327	0.063	0.037	0.049	0.585	0.467	1.858	0.745	0.214	0.045	0.024	0.044
RMSE	0.935	0.790	2.814	1.653	0.342	0.068	0.041	0.066	0.605	0.503	1.870	1.234	0.227	0.054	0.028	0.051

	$T = 200$								$T = 500$							
	a_2	a_4	c_1	b_2	d_1	σ_1	σ_2	ρ	a_2	a_2	c_1	b_2	d_1	σ_1	σ_2	ρ
Parameter set 3																
True	1.5	-3.0	-10	-1.5	10	1.0	0.5	-0.9	1.5	-3.0	-10	-1.5	10	1.0	0.5	-0.9
Mean	2.374	-3.289	-9.944	-2.018	10.576	0.949	0.503	-0.825	2.429	-3.346	-9.917	-1.953	10.532	0.960	0.507	-0.796
FSSE	0.880	1.010	0.589	0.470	0.495	0.068	0.055	0.141	0.652	0.764	0.396	0.346	0.302	0.058	0.049	0.119
RMSE	1.239	1.048	0.590	1.121	0.759	0.084	0.055	0.160	1.134	0.837	0.403	1.013	0.612	0.070	0.050	0.158
True	1.5	-3.0	-10	-1.5	10	1.0	0.5	-0.5	1.5	-3.0	-10	-1.5	10	1.0	0.5	-0.5
Mean	1.986	-3.245	-9.823	-1.695	10.279	0.971	0.504	-0.437	1.893	-3.087	-9.896	-1.690	10.218	0.971	0.520	-0.392
FSSE	0.927	1.094	0.647	0.350	0.477	0.062	0.058	0.154	0.579	0.713	0.356	0.213	0.260	0.037	0.040	0.101
RMSE	1.044	1.118	0.669	0.778	0.551	0.068	0.058	0.166	0.699	0.717	0.370	0.722	0.338	0.047	0.044	0.148
True	1.5	-3.0	-10	-1.5	10	1.0	0.5	0	1.5	-3.0	-10	-1.5	10	1.0	0.5	0
Mean	1.766	-3.332	-9.833	-1.469	10.214	0.980	0.495	0.011	1.625	-3.081	-9.851	-1.494	10.153	0.981	0.511	0.022
FSSE	0.920	1.211	0.508	0.327	0.355	0.062	0.057	0.122	0.596	0.739	0.567	0.225	0.394	0.048	0.049	0.089
RMSE	0.956	1.253	0.534	0.571	0.413	0.065	0.057	0.122	0.607	0.741	0.585	0.543	0.422	0.051	0.050	0.092
True	1.5	-3.0	-10	-1.5	10	1.0	0.5	0.5	1.5	-3.0	-10	-1.5	10	1.0	0.5	0.5
Mean	1.416	-3.272	-9.787	-1.279	10.312	0.980	0.482	0.482	1.334	-3.024	-9.810	-1.321	10.287	0.980	0.497	0.453
FSSE	0.971	1.368	0.528	0.360	0.357	0.066	0.058	0.144	0.567	0.785	0.336	0.262	0.263	0.037	0.043	0.087
RMSE	0.972	1.391	0.568	0.454	0.474	0.069	0.061	0.145	0.590	0.784	0.385	0.414	0.389	0.042	0.043	0.099
True	1.5	-3.0	-10	-1.5	10	1.0	0.5	0.9	1.5	-3.0	-10	-1.5	10	1.0	0.5	0.9
Mean	0.535	-2.449	-9.696	-1.114	10.665	0.946	0.460	0.932	0.389	-2.345	-9.711	-1.086	10.664	0.968	0.469	0.924
FSSE	1.056	1.322	0.622	0.392	0.402	0.062	0.060	0.085	0.708	0.870	0.452	0.289	0.303	0.048	0.046	0.072
RMSE	1.429	1.429	0.691	0.407	0.776	0.083	0.072	0.090	1.316	1.087	0.535	0.301	0.729	0.058	0.056	0.076

Note: The table reports the statistics of 200 Monte Carlo runs of each parameter set with sample sizes $T = 200$ and $T = 500$. Discrete observations have been extracted at time intervals $\Delta t = 1/12$. For the numerical approximations to the likelihood function: the ADI method of McKee *et al.* has been used with grid increments $\kappa = 1/144$ for the time dimension. For the space dimensions the number of grid points has been kept fixed at $N_x = N_y = 50$, but for the the grid borders some flexibility has been allowed: a grid symmetric about 0 has been chosen with x_{max} and y_{max} the next higher multiples of 0.1 larger than the maximum of the series of absolute observations of a simulation and $x_{min} = -x_{max}$, $y_{min} = -y_{max}$. Note that the time scale of the discretely sampled data (with $\Delta t = 1/12$ rather than $\Delta t = 1$ as in the bivariate and trivariate diffusions of Tables 1 and 3) is determined by the intrinsic time scale of the drift terms.

Table 3: Monte Carlo Results for Trivariate Diffusion

	$T = 200$						$T = 500$					
	σ_1	σ_2	σ_3	ρ_{12}	ρ_{23}	ρ_{13}	σ_1	σ_2	σ_3	ρ_{12}	ρ_{23}	ρ_{13}
Parameter set 1												
$True$	0.2	0.1	0.2	0	-0.5	-0.5	0.2	0.1	0.2	0	-0.5	-0.5
EML												
$Means$	0.199	0.099	0.199	0.001	-0.486	-0.488	0.200	0.099	0.199	0.003	-0.490	-0.491
$FSSE$	0.011	0.005	0.010	0.069	0.058	0.067	0.008	0.003	0.007	0.047	0.040	0.041
$RMSE$	0.011	0.005	0.010	0.069	0.060	0.067	0.008	0.003	0.007	0.047	0.041	0.042
$FP-ML$												
$Means$	0.200	0.100	0.228	0.003	-0.446	-0.445	0.201	0.100	0.229	0.006	-0.448	-0.447
$FSSE$	0.011	0.005	0.110	0.073	0.077	0.071	0.009	0.004	0.108	0.056	0.067	0.061
$RMSE$	0.011	0.005	0.113	0.073	0.094	0.089	0.009	0.004	0.112	0.056	0.085	0.080
Parameter set 2												
$True$	0.2	0.1	0.2	0.5	0	-0.5	0.2	0.1	0.2	0.5	0	-0.5
EML												
$Means$	0.198	0.099	0.200	0.489	-0.001	-0.485	0.200	0.099	0.199	0.495	-0.004	-0.490
$FSSE$	0.011	0.005	0.012	0.062	0.060	0.067	0.007	0.004	0.007	0.046	0.044	0.046
$RMSE$	0.011	0.005	0.012	0.063	0.060	0.069	0.007	0.004	0.007	0.046	0.044	0.047
$FP-ML$												
$Means$	0.200	0.100	0.229	0.484	0.003	-0.439	0.200	0.100	0.228	0.488	-0.005	-0.442
$FSSE$	0.011	0.005	0.109	0.050	0.061	0.077	0.008	0.003	0.110	0.034	0.045	0.064
$RMSE$	0.011	0.005	0.112	0.053	0.061	0.098	0.008	0.003	0.113	0.036	0.046	0.086

	$T = 200$						$T = 500$					
	σ_1	σ_2	σ_3	ρ_{12}	ρ_{23}	ρ_{13}	σ_1	σ_2	σ_3	ρ_{12}	ρ_{23}	ρ_{13}
Parameter set 3												
$True$	0.2	0.1	0.2	0.9	-0.9	-0.9	0.2	0.1	0.2	0.9	-0.9	-0.9
EML												
$Means$	0.199	0.099	0.200	0.893	-0.892	-0.883	0.200	0.100	0.200	0.894	-0.893	-0.886
$FSSE$	0.011	0.006	0.011	0.052	0.052	0.115	0.007	0.005	0.007	0.056	0.056	0.112
$RMSE$	0.011	0.006	0.011	0.052	0.053	0.116	0.007	0.005	0.006	0.056	0.056	0.112
$FP-ML$												
$Means$	0.197	0.099	0.207	0.893	-0.856	-0.855	0.198	0.099	0.214	0.892	-0.854	-0.854
$FSSE$	0.010	0.005	0.018	0.016	0.066	0.066	0.007	0.004	0.071	0.013	0.068	0.066
$RMSE$	0.011	0.005	0.019	0.017	0.079	0.079	0.007	0.004	0.072	0.015	0.082	0.080
Parameter set 4												
$True$	0.2	0.1	0.2	0.25	0.5	-0.5	0.2	0.1	0.2	0.25	0.5	-0.5
EML												
$Means$	0.197	0.098	0.197	0.232	0.471	-0.472	0.199	0.099	0.198	0.242	0.478	-0.481
$FSSE$	0.011	0.006	0.012	0.071	0.078	0.079	0.008	0.004	0.009	0.053	0.062	0.059
$RMSE$	0.012	0.006	0.012	0.073	0.083	0.083	0.008	0.004	0.009	0.054	0.065	0.062
$FP-ML$												
$Means$	0.198	0.099	0.219	0.240	0.440	-0.437	0.199	0.099	0.219	0.243	0.436	-0.439
$FSSE$	0.011	0.005	0.029	0.061	0.062	0.075	0.006	0.003	0.028	0.041	0.056	0.058
$RMSE$	0.011	0.005	0.035	0.061	0.086	0.098	0.006	0.003	0.034	0.042	0.085	0.084

	$T = 200$						$T = 500$					
	σ_1	σ_2	σ_3	ρ_{12}	ρ_{23}	ρ_{13}	σ_1	σ_2	σ_3	ρ_{12}	ρ_{23}	ρ_{13}
Parameter set 5												
True	0.2	0.1	0.2	0.95	0	0	0.2	0.1	0.2	0.95	0	0
EML												
Means	0.199	0.099	0.200	0.948	-0.014	-0.016	0.197	0.098	0.202	0.944	-0.002	-0.007
FSSE	0.012	0.007	0.010	0.010	0.045	0.047	0.012	0.006	0.013	0.022	0.037	0.035
RMSE	0.012	0.007	0.010	0.010	0.047	0.049	0.013	0.006	0.013	0.023	0.037	0.035
$FP-ML$												
Means	0.195	0.097	0.223	0.954	-0.010	-0.010	0.196	0.098	0.222	0.946	-0.002	-0.002
FSSE	0.010	0.005	0.046	0.020	0.056	0.056	0.007	0.003	0.034	0.018	0.041	0.040
RMSE	0.011	0.006	0.051	0.020	0.056	0.057	0.008	0.004	0.040	0.019	0.041	0.040

Note: The table reports the statistics of 200 Monte Carlo runs of each parameter set with sample sizes $T = 200$ and $T = 500$. Discrete observations have been extracted at time intervals $\Delta t = 1$. For the numerical approximations to the likelihood function the ADI method of McKee *et al.* has been used with grid increments $\kappa = 1/12$ for the time dimension. For the space dimensions the number of grid points has been kept fixed at $N_{x_1} = N_{x_2} = N_{x_3} = 50$, but for the grid borders some flexibility has been allowed: a grid symmetric about 0 has been chosen with $x_{1,max}$, $x_{2,max}$ and $x_{3,max}$ the next higher multiples of 0.1 larger than the maximum of the series of absolute observations of a simulation.

Table 4: Parameter Estimates for Bivariate Model of Medium-Run Sentiment and Prices

M-Sent and Prices			
Param.	Model I	Model II	Model III
b_1	0.004 (0.017)		
b_2	-0.331 (0.157)	-0.316 (0.121)	-0.099 (0.055)
b_3	0.029 (0.905)		
b_4	-0.009 (2.774)		
d_2	0.001 (0.000)	0.001 (0.000)	-0.001 (0.000)
e_0	17.072 (12.983)	17.075 (8.409)	13.164 (11.265)
e_2	174.413 (123.409)	174.413 (105.237)	175.340 (101.045)
σ_y	0.126 (0.017)	0.127 (0.015)	0.082 (0.006)
σ_p	110.490 (5.010)	110.490 (4.768)	105.029 (5.256)
ρ_{yp}	-0.785 (0.091)	-0.797 (0.073)	
lkl	-1109.042	-1109.163	-1115.610
AIC	2238.084	2232.327	2243.220
BIC	2268.124	2253.355	2261.243

Note: The model has been estimated via numerical approximation of the transient density with a time increment of $\kappa = 1/12$ and $N_y = N_p = 100$ equidistant space increments in both dimensions. Standard errors of estimated parameters are given in parentheses.

Table 5: Parameter Estimates for Bivariate Model of Short-Run Sentiment and Prices

S-Sent and Prices			
Param.	Model I	Model II	Model III
a_1	0.091 (0.045)	0.106 (0.052)	0.057 (0.025)
a_2	-0.328 (0.168)		
a_3	-0.093 (0.176)		
a_4	-1.595 (0.538)	-2.260 (0.290)	-3.017 (0.271)
c_2	0.006 (0.001)	0.002 (0.001)	0.007 (0.000)
e_0	34.519 (14.622)	27.188 (12.252)	26.320 (11.162)
e_1	-44.070 (49.470)		
σ_x	0.127 (0.058)	0.354 (0.047)	0.114 (0.035)
σ_p	110.646 (6.620)	106.088 (5.014)	104.291 (4.731)
ρ_{xp}	0.512 (0.178)	0.827 (0.069)	
lkl	-1122.277	-1122.117	-1126.326
AIC	2264.554	2258.234	2264.652
BIC	2294.593	2279.261	2282.675

Note: The model has been estimated via numerical approximation of the transient density with a time increment of $\kappa = 1/12$ and $N_x = N_p = 100$ equidistant space increments in both dimensions. Standard errors of estimated parameters are given in parentheses.

Table 6: Parameter Estimates for Interaction of Short-Run and Medium-Run Sentiment

S-Sent and M-Sent			
Param.	Model I	Model II	Model III
a_1	0.162 (0.131)	0.132 (0.076)	0.176 (0.086)
a_2	1.249 (0.963)	1.284 (0.944)	1.285 (0.896)
a_3	-0.375 (0.890)		
a_4	-7.399 (4.391)	-7.877 (4.209)	-7.751 (3.999)
c_1	1.331 (0.644)	1.308 (0.621)	0.579 (0.470)
b_1	0.017 (0.016)	0.017 (0.011)	0.017 (0.011)
b_2	-0.294 (0.109)	-0.236 (0.068)	-0.233 (0.065)
b_3	0.173 (1.040)		
b_4	0.427 (2.603)		
d_1	-0.011 (0.045)		
σ_x	0.522 (0.099)	0.524 (0.089)	0.513 (0.086)
σ_y	0.095 (0.008)	0.095 (0.007)	0.095 (0.007)
ρ_{xy}	-0.440 (0.097)	-0.444 (0.080)	
lkl	-1152.817	-1153.457	-1162.595
AIC	2331.633	2324.915	2341.191
BIC	2370.685	2351.950	2365.222

Note: The model has been estimated via numerical approximation of the transient density with a time increment of $\kappa = 1/12$ and $N_x = N_y = 100$ equidistant space increments in both dimensions. Standard errors of estimated parameters are given in parentheses.

Table 7: Parameter Estimates for Trivariate Model of Short-Run Sentiment, Medium-Run Sentiment and Prices

S-Sent, M-Sent and Prices				
Param.	Model I	Model II	Model III	Model IV
a_4	-2.605 (0.290)	-2.959 (0.281)	-2.736 (0.298)	-2.725 ()
c_1	0.869 (0.345)	0.723 (0.216)	0.375 (0.175)	0.407 ()
c_2	0.003 (0.001)	0.003 (0.001)	0.006 (0.001)	0.007 ()
b_1	0.024 (0.011)	0.018 (0.010)	0.024 (0.010)	0.021 ()
b_2	-0.248 (0.079)	-0.257 (0.072)	-0.250 (0.068)	-0.225 ()
e_2	236.772 (84.296)	138.077 (72.894)	223.465 (67.991)	203.201 ()
σ_x	0.317 (0.041)	0.288 (0.036)	0.183 (0.043)	0.047 ()
σ_y	0.099 (0.009)	0.098 (0.0070)	0.096 (0.007)	0.092 ()
σ_p	107.102 (5.602)	106.311 (6.332)	104.198 (6.068)	102.984 ()
ρ_{xy}	-0.309 (0.062)	-0.040 (0.065)		-0.195 ()
ρ_{yp}	-0.416 (0.058)		-0.342 (0.071)	-0.378 ()
ρ_{xp}	0.769 (0.074)	0.797 (0.053)	0.267 (0.198)	
lkl	-1196.146	-1369.984	-1266.590	-1327.501
AIC	2416.293	2761.968	2555.180	2677.001
BIC	2452.340	2795.012	2588.224	2710.045

Note: The model has been estimated via numerical estimation of the transient density with a time increment of $\kappa = 1/8$ and $N_x = N_y = N_p = 50$ equidistant space increments in all three dimensions. Standard errors of estimated parameters are given in parentheses. Since the covariance matrix of the parameters failed to invert, no standard errors could be obtained for model IV.

Table 8: RMSEs of Out-of-Sample Forecasts from Trivariate Model

horiz.	single period	cumulative returns
1	1.003	1.003
2	0.993	0.993
3	0.997	0.991*
4	0.988*	0.973*
5	0.992*	0.960**
6	0.994	0.955**
7	0.990*	0.945**
8	0.993*	0.939**

Note: The table shows relative root mean-squared errors (RMSEs) of single period and cumulative forecasts of DAX returns from the trivariate model (i.e. model I of Table 7) for single and cumulative forecasts over horizons ranging from one to eight weeks. Relative RMSEs have been obtained by dividing the original RMSE by that of Brownian motion with drift as benchmark. * and ** denote significantly better predictive capability of the trivariate diffusion against the Brownian benchmark at the 95 and 99 percent significance level, respectively, under the Diebold-Mariano test for forecast comparison of nested models.