An Inverse Problem for a Nonlinear Stochastic Differential Equation Modeling Price Dynamics

SIMON JÄGER[1] and EKATERINA KOSTINA[2]

[1] *Department of Economics, University of Bonn, Adenauerallee 24, D-53113 Bonn, Germany, Email: simon.jaeger@uni-bonn.de*

[2] *Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany, Email: ekaterina.kostina@iwr.uni-heidelberg.de*

Abstract

Diffusion processes are widely used for mathematical modeling in finance e.g. in modeling foreign exchange rates. This paper presents a non-linear stochastic continuous-time model that captures the main characteristics of price dynamics. The generalized mean reversion process discloses various features of observed price movements such as multi-modality of the distributions, multiple equilibria, and regime switching. The attractors depend substantially on the economic environment. The model reveals a significant connection between exchange rates and its fundamentals. Furthermore, it is consistent with traditional flexible exchange rate models.

Stochastic differential equations describing diffusion processes are directly linked to the forward Kolmogorov equation. In order to calibrate the models, efficient algorithms identifying the system parameters are in demand. Taking into account nonlinear effects in volatility and drift and dependence on observed economic data, which are not directly modeled, one obtains problems which cannot be treated by standard numerical methods. The coefficients are rapidly oscillatory and strong instabilities may arise. To handle these problems we develop numerical methods, which are used to simulate the nonlinear dynamics of exchange rates depending on economic data.

Keywords: generalized mean reversion, multiple equilibria, numerical methods for inverse problems, forward Kolmogorov equation, Gauss-Newton

1 Introduction

Modeling

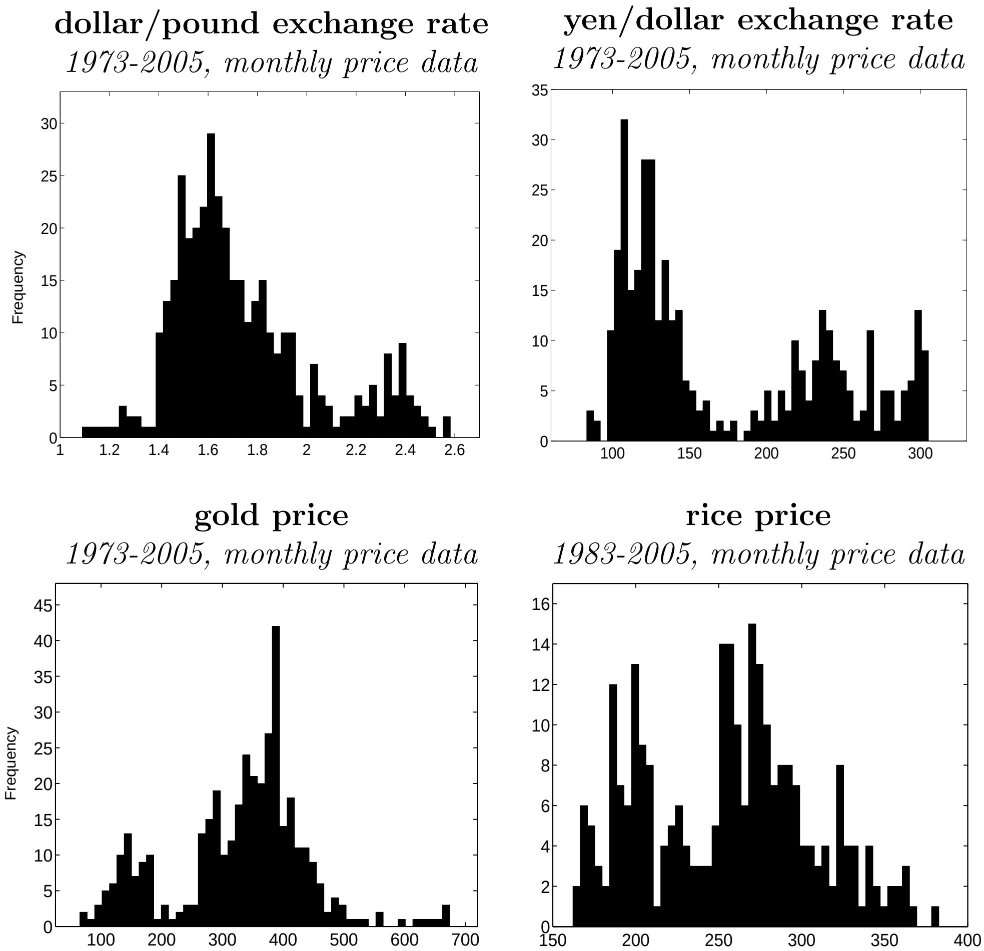
Mathematical modeling and simulation have become important tools for the analysis of data and the prediction of economic and financial processes. For a long time, mainly stationary systems and stationary fixpoints of model equations were considered, not taking properly into account the importance of dynamical effects. Therefore, after thirty years of flexible price movements, modeling the development of foreign exchange rates remains a challenge. Beginning with the seminal empirical work of Meese et al. (1983) it still seems questionable whether any structural exchange rate model would be of systematic value. Usually, a pure random walk process outperforms all classical exchange rate approaches based on monetary fundamentals. Hence one is inclined to conclude that flexible exchange rates are ordinary stochastic processes in which the different states of the economic environment are of secondary importance.

Another defect is getting more and more obvious in the analysis of economic data sets: There is a lack of combining stochastics and nonlinear dynamical systems in methodology, pure statistics and modeling based on economic facts in theory. Model based statistics has to be developed in order to integrate and exploit economic knowledge better. Stochastic nonlinear dynamical systems, describing the arising processes more adequate, have to be investigated with the aim to get better qualitative and more precise quantitative answers.

Standard approaches in modeling price dynamics e.g. foreign exchange rates and commodity prices are working with the hypothesis of a single long run price equilibrium (see e.g. Creedy et al. (1996) and Geman (2005)). Deviations from this reversion level are expected to be temporary, thus the dynamics is mainly driven by one attracting equilibrium. However, the interplay of nonlinearities in the dynamics and the stochastic influences in the system are highly important, but not enough taken into account. These interactions may lead to effects which cannot be explained otherwise: e.g. multi-modal distributions can be traced back to multiple states in the dynamical system, observed jumps and strong oscillation in the historical data can be explained by stochastic changes of attractors. Small random perturbations may push a balanced market from one equilibrium into another, reflecting both regime switches and rare events.

Nonlinear phenomena in real data can be observed in different asset markets and are getting more and more into focus of mathematical modeling of financial price processes. Recent analysis of spot and futures prices (e.g. by Borovkova (2003)) detect a similar behavior in case of agricultural and energy commodities. The histograms in Figure 1 illustrates the clustering of different exchange rates, gold, and rice prices.

Fig. 1. Price histograms



Source: International Monetary Fund.

Parameter estimation

An important issue in finance is the parameter estimation (inverse problem) of stochastic differential equations. Recent approaches for modeling the dynamics of asset prices such as interest rates, commodity prices, or foreign exchange rates are based on diffusion processes with nonlinear drift terms and nonlinear volatility functions (see e.g. Ait-Sahalia (1999)). In order to apply the model to predictions or control of real processes, the model has to be able to reproduce the real process data quantitatively under changing conditions. Values for the unknown parameters and the initial data of the initial value problem have to be estimated, such that the observed behavior of the economic process is reproduced in an optimal way. We estimate the parameters without the assumption of stationary distributions. To make sure to identify the system parameter considering the full time dependent situation. Thus, we generalize results obtained for the quasi steady state situation.

We provide a general framework to estimate unknown parameters of price diffusion processes used in mathematical finance. The principle idea is based on minimizing a weighted least squares functional constrained by the forward Kolmogorov equation capturing the dynamics of the price probability density distribution. Since there already exists an extensive research on solving inverse problems for diffusion-transport equations, it seemed rather natural to use the forward Kolmogorov equation to determine the missing parameters. However, it soon became obvious, that the coefficients are rapidly oscillatory and strong instabilities may arise such that the available algorithms did not work well enough. Therefore, it was necessary to improve and to adjust the numerical methods for the inverse problem.

2 Nonlinear price dynamics

It is common practice in financial modeling that the price dynamics X is modeled by an Itô stochastic differential equation:

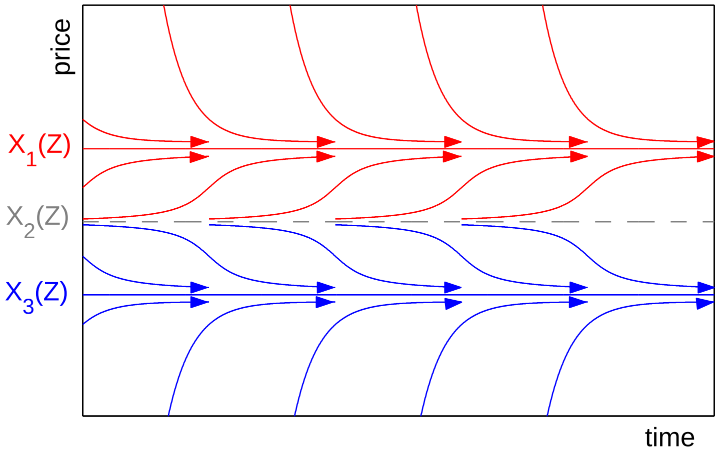
$$dX = \mu(t, X(t), Z(t))dt + \sigma(t, X(t), Z(t))dW. \tag{1}$$

Here $Z(t)$ are external, such as economic or political effects and W is a standard Wiener process with the property that dW is distributed as $\mathcal{N}(0, dt)$, and μ and σ satisfy Lipschitz and growth conditions sufficient for the existence of a continuous solution to (1). In case that all necessary coefficient of the model equation are known, solutions to the stochastic differential equations can be computed using available algorithms (compare e.g. Kloeden et al. (1992)). One obtains a realization of the trajectory of the system. However, in reality the drift term $\mu(\cdot)$ and the volatility $\sigma(\cdot)$ are unknown and need to be determined by modeling and from data by solving the inverse problem. Since there exists no theoretical approach determining the unknown parameters, it is necessary to identify the model and its parameters using observable data of the real process.

2.1 Generalized mean reversion process

We present a nonlinear mean reversion process capturing main characteristics of price dynamics. The nonlinearity is chosen such that canceling the noise term leads to a deterministic dynamics which is characterized by the evolution of so-called quasi-steady states. These vary in time and play the role of steady states in an autonomous system. They are called "stable" or "unstable" states if they are locally in time attractive or repulsive. It is important

Fig. 2. Generalized mean reversion process, two stable equilibria



This figure illustrates the root decomposition of the price dynamics. In case of constant equilibria the quasi-steady state X_2 is repelling, whereas X_1 and X_3 are attracting. By crossing X_2 due to e.g. small random perturbations, the domain of attraction is changed and a new attracting level might be reached. However, we expect these quasi-steady states X_j to be dependent on key variables such as interest rates at home and abroad. As a consequence, the attractors may change both their characteristics and location.

that time intervals may arise with several quasi-steady states. In these intervals stochasticity can lead to transitions between these time-varying states, to jumps and a behavior similar to instabilities. We take the simplest nonlinearity allowing such a behavior: a polynomial of order three. Thus, the nonlinear drift function can be written as a product of linear factors:

$$\mu(t, X, Z) = C_0(X_1(Z) - X)(X_2(Z) - X)(X_3(Z) - X), \tag{2}$$

where $X_1(Z)$, $X_2(Z)$ and $X_3(Z)$ are the roots of the polynomial, Z a set of economic variables, and C_0 the speed of adjustment coefficient, $C_0 > 0$. Without loss of generality, X_1 is real. If X_2 is complex and not real, $X_3 = \overline{X}_2$, e.g. the drift term has only one real root. This root decomposition demonstrates the underlying dynamics behind the price evolution. Similar to the mean reversion process, which is just a special case of our model, the price is pulled back to some long-run price level at a rate C_0. In absence of randomness and constant long-run equilibria the process reverts to its gravitational cores X_1 or X_3 (see Figure 2)

In the following the attractors evolve in time changing their location and possibly their stability, thus creating natural zone of instabilities.

2.2 Quasi steady states

In economics, there is rather often assumed that the trajectories of the system are tending to an equilibrium for large times. However, in reality this assumption is not valid, also the systems will not be autonomous. The roots X_j of $\mu(t, X(t), Z(t))$ will depend on time t and a fundamental economic variables $Z(t)$. If changes in time are slow, it is not unrealistic to assume that $dX_j = dt$ is very small, that means that one $X_j(t)$ play locally in time the role of equilibria. We call them quasi-steady-states. Accordingly, the quasi stationary cases are given by $X = X_j(Z)$ for all $j = 1, 2, 3$.

2.3 Models of exchange rate determinants

In order to determine the dynamics of our steady-states, we refer to monetary exchange rate models which go back to Frenkel (1976), Mussa (1976), and Bilson (1978). It is widely accepted, that these initial exchange rate models have some validity when considered as a long-run equilibrium (see e.g. MacDonald et al. (1995)). The principle idea is that the foreign exchange rate dynamics is driven by the relative behavior of a set of underlying economic variables (home versus foreign variables). The fundamental monetary price equation can be summarized as

$$x = \alpha_1 m + \alpha_2 m^* + \alpha_3 y + \alpha_4 y^* + \alpha_5 i + \alpha_6 i^*, \tag{3}$$

where m is the logarithm of the money supply, y is the logarithm of real income, i is the nominal interest rate, and x is the logarithm of the exchange rate. Variables with an asterisk denotes determinants that correspond to the foreign country. In the literature there is a controversial discussion about which economic variables should be included and even the direction of influence is ambiguous. In consequence, the traditional flexible exchange rate approach only serve as a common reference point. A selective literature survey on the economics of exchange rates over the last decades is offered e.g. by Taylor (1995).

2.4 Modeling the dynamics of attractors

By modeling the dynamics of the quasi-steady states as a product of N economic key variables (e.g. money supply M, real output Y, and the exponential of nominal interest rates i)

$$X_j(Z) = C_j Z_1^{\alpha_{j1}} \cdots Z_N^{\alpha_{jN}}, \quad j = 1, 2, 3, \tag{4}$$

the standard linear monetary exchange rate model is embedded into the nonlinear model (compare equation (3)). Hence, the dynamics of the attractors and with that of the exchange rates are driven by well known economic relationships.

As first step, we choose a constant volatility. Further investigation could be done by taking into account the influence of economic fundamentals and lagged effects on the volatility. However, the estimation results seem to justify the concentration at the time being on modeling the drift term.

3 Parameter estimation

3.1 Problem formulation

In order to solve the inverse problem for stochastic differential equations modeling the relevant processes, this paper is solving the inverse problem of the corresponding forward Kolmogorov equation. This is a non stochastic partial differential containing a diffusion and a drift term. The parameters of these terms have to be recovered numerically from the available data. Since there exists already an extensive research on solving inverse problems for diffusion-transport equations, it seemed to be rather natural to use the forward Kolmogorov equation to determine the missing parameters. However, it very soon became obvious, that the arising coefficients are such that the available algorithms did not work well enough. Therefore, it was necessary to improve and to adjust the numerical methods for the inverse problem. This paper is presenting and testing an improved algorithm overcoming these difficulties.

Traditionally, given the probability density distribution of the price process, the maximum likelihood method can be applied to estimate the unknown parameters. Except for very simple linear drift and volatility functions, the forward Kolmogorov equation cannot be solved explicitly and the likelihood function cannot be given in an analytic formula. However, this is possible e.g. for stationary situation in case of one space dimension. Creedy et al. (1996) applied this idea to estimate nonlinear exchange rate dynamics. By contrast, this paper is considering the full time dependent situation and, thus, generalizes results obtained for the quasi steady state situation. Beginning with the formulation of the estimation problem we describe a generalized Gauss-Newton algorithm constrained by the forward Kolmogorov equation and some initial and boundary conditions. Therefore, we do not assume the restrictive hypothesis of an time-independent distribution and, thus, generalize the equilibrium approach.

3.2 Distributional dynamics

We consider a stochastic process $X_t, t \in [t_0, T]$ with the probability space $(\Omega, \mathcal{F}, P)$ and the distribution $F_t(x) = P(X_t \leq x), t \in [t_0, T]$ and $x \in \mathbb{R}$. It is assumed that the drift term and the volatility function of the stochastic process (1) depend on the spatial variable x and the unknown parameter vector p: $\mu = \mu(t, x; p)$ and $\sigma^2 = \sigma^2(t, x; p)$. To estimate the unknown parameters we make use of the forward Kolmogorov equation for the density function $f(t, x) = \frac{dF_t(x)}{dx}$ of the stochastic process X_t. Stochastic differential equations describing the diffusion processes can be directly linked to the forward Kolmogorov equation. A detailed derivation of the link between (1) and (5) is offered by e.g. Stroock and Varadhan (1979). The transitional distribution at each point of time satisfies the partial differential equation:

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}(\mu f) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(\sigma^2 f). \tag{5}$$

In some special cases this equation, also known as Fokker-Planck equation can be explicitly solved. However, this is no longer true for complex price dynamics. For one state variable x, stationary solutions to (5) can be simply computed by direct integration.

3.3 Stationary approach

The stationary density f^* of a nonlinear diffusion process is found by setting $\partial f/\partial t = 0$. This converts the diffusion equation (5) for $n = 1$ into an ordinary differential equation for the stationary density, which can be determined by integration as

$$f^*(x) = \exp\left(-\int_0^x \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi - 2\ln\sigma(x)\right)\eta^\star. \tag{6}$$

The normalizing constant $\eta^\star$ is chosen such that the integral of f^* over its domain is 1. It is nothing but straightforward to use the stationary distribution (6) to estimate the unknown parameter via the well known maximum-likelihood technique. Creedy et al. (1996), for instance, make use of this idea to identify the drift term and volatility function of a nonlinear exchange rate model.

In general, quasi stationarity is assuming that the trajectories of the underlying process are tending very fast to stationary points. Creedy et al. (1996): *"if prices are flexible, the speed of convergence to the stationary distribution is* *fast"*. This assumption may not be justified in real situations. In fact, the existence of simple stationary points by itself cannot be assumed in reality, since the model systems are in general not autonomous. There will exist states which evolve slowly in time and locally play the role of stationary points. In the following we use the expression stationary states also for such states, despite the fact they may evolve in time slowly. The restriction to stationary distributions in strict sense has to be considered as an approximation, which can be too rigorous. Taking into account real time dependent data, as also the authors just mentioned are doing, one should consider the full time-dependence.

3.4 Non-stationary approach

In the following, we drop the assumption of a stationary distribution. Using the fact that the price distribution of the stochastic price process satisfies the forward Kolmogorov equation, we estimate the unknown parameters p, by solving the following optimization problem

Minimize the weighted least squares functional

$$\min_p \sum_{j=1}^{N_{meas}} \left(\eta_j - \int_0^\infty f(t_j, x, p) x dx \right)^2 / \omega_j^2, \tag{7}$$

subject to the forward Kolmogorov equation

$$\frac{\partial f(t,x,p)}{\partial t} = -\frac{\partial(\mu(t,x,p) f(t,x,p))}{\partial x} + \frac{1}{2}\frac{\partial^2(\sigma^2(t,x,p) f(t,x,p))}{\partial x^2}, \quad 0 \le t \le T, \tag{8}$$

a state condition

$$\int_0^\infty f(t,x,p) dx = 1, \quad 0 \le t \le T, \tag{9}$$

an initial condition

$$f(t,x,p) \mid_{t=0} = f_0(x,p), \tag{10}$$

and two boundary conditions

$$\mu(t,x,p) - \frac{1}{2}\frac{\partial(\sigma^2(t,x,p) f(t,x,p))}{\partial x}\bigg|_{x=x_{min}} = 0, \tag{11}$$

$$\mu(t,x,p) - \frac{1}{2}\frac{\partial(\sigma^2(t,x,p)f(t,x,p))}{\partial x}\bigg|_{x=x_{max}} = 0, \quad 0 \le t \le T \ .$$

Here, the least squares functional (7) can be interpreted as a weighted norm of the difference between the real values η_j of the random variable x at time points $t_j, j = 1, ..., N_{meas}$, and their expected values. The parameters p will be estimated by minimizing this functional subject to the forward Kolmogorov equation for the density function $f(t,x,p)$ (8), the state condition (9), initial conditions (10) and boundary conditions (11). In the paper, we assume that the initial density $f_0(x,p)$ is given by

$$f_0(x,p) := \exp\left(-(x-x_0)^2\right)\eta^\star, \quad \eta^\star \text{ is a normalizing constant,}$$

with an additional parameter x_0 to estimate.

4 Numerical methods

The optimization problem is a parameter estimation problem with partial differential equations as constraints. To solve this problem we apply the so-called boundary value problem (BVP) approach, see Bock (1987). The basic idea consists in modeling the dynamic equations like a boundary value problem and then performing simultaneously the minimization of the cost function subject to the constraints given by the discretized boundary value problem. We apply a generalized Gauss-Newton methods with trust-region globalization techniques to solve the nonlinear least squares problem using a tailored linear algebra to exploit the special structures arising from the multiple shooting discretization. In the following subsections we illustrate the basic ideas of the applied numerical methods.

4.1 Spatial discretization

In a first step, we reduce the forward Kolmogorov equation (8) into a system of ordinary differential equations (ODE) employing the method of lines Schiesser (1991). The initial conditions (10) and the the integrals appearing in the problem formulation (7)-(11) are transformed accordingly. The state condition (9) and the boundary conditions (11) are taken into account in the transformation of the forward Kolmogorov equation (8) and are implicitly included into the resulting ODE system. As a consequence, the parameter estimation problem (7)-(11) results in a nonlinear least squares ODE constrained problem which can be formally written as

$$\min_p \; ||r_1(y,p)||_2^2 := \sum_{j=1}^{m_1} \left(\eta_j - B(t_j, y, p)\right)^2 / \omega_j^2 \qquad (12)$$
$$\text{s.t. } \dot{y} = \phi(t, y, p), \quad 0 \le t \le T, \text{ and } y(0) = y_0.$$

For solving problem (12) we use the boundary value problem approach according to which the ODEs are parameterized by multiple shooting and are treated as implicit constraints in the minimization problem.

4.2 Parameterization in time - Multiple shooting

We parameterize the semidiscretized parameter estimation problem (12) in time using multiple shooting approach. The scheme of the multiple shooting consists in the following. First, one chooses a suitable grid of multiple shooting nodes τ_j

$$0 = \tau_0 < \tau_1 < \ldots < \tau_m = T,$$

covering the interval where measurements are given.

At each grid point the values of the state variables y_j are chosen as additional unknowns and m initial value problems

$$\dot{y} = \phi(t, y, p), \quad y(\tau_j) = y_j, \qquad (13)$$

are solved on each subinterval $I_j := [\tau_j, \tau_{j+1}]$ to yield a solution $y(t; y_j, p)$ for $t \in I_j$. The principle of multiple shooting is depicted in the figure below.

Solutions of dynamic systems, generated by this procedure, are usually not continuous at τ_j. This has to be enforced by additional matching conditions

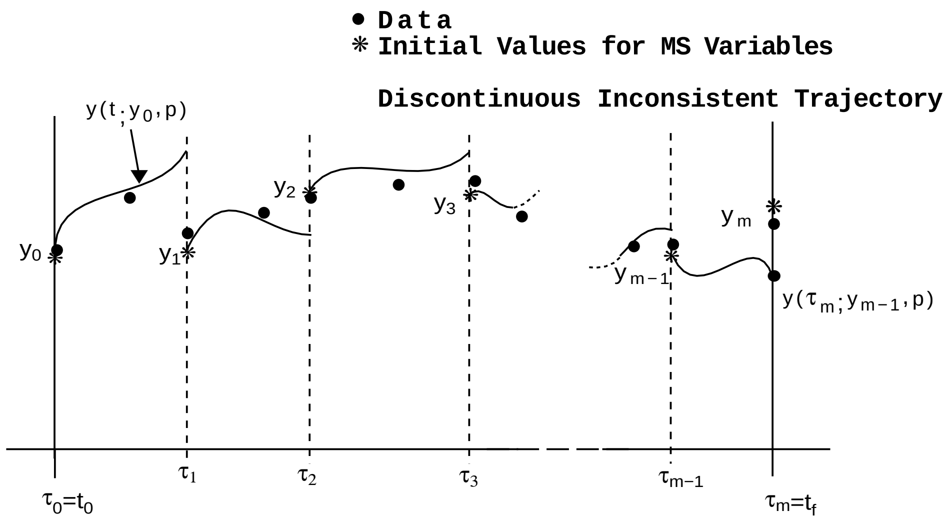
$$h_j(y_j, y_{j+1}, p) := y(\tau_{j+1}; y_j, p) - y_{j+1}, \quad j = 0, ..., m-1.$$

Inserting the computed values $y(t_i, y_j, p)$, $\tau_j \le t_i \le \tau_{j+1}$, into problem (12) one obtains a constrained problem in the variables $(y, p) := (y_0, \ldots, y_m, p)$:

$$\min \; ||r_1(y,p)||_2^2 \qquad (14)$$
$$\text{s.t. } h_j(y_j, y_{j+1}, p) = 0, \quad j = 0, ..., m-1.$$

Multiple shooting possesses several advantages which are discussed to large extent e.g. in Bock (1987).

Fig. 3. Multiple Shooting



4.3 Generalized Gauss-Newton method with trust-region globalization

For the solution of nonlinear constrained least squares problems of the presented type, Bock (1983) proposed a generalization of the Gauss-Newton Method which was only applicable to unconstrained least squares problems. The numerical method has proven to be stable and efficient for a series of real life parameter estimation problems constrained by ordinary differential equations and differential-algebraic equations.

The parameterization of the dynamic system yields to a finite dimensional, possibly large scale, nonlinear equality constrained approximation problem, which can be formally written as

$$\begin{aligned} &\min \, ||r_1(s)||_2^2, \\ &\text{s.t. } r_2(s) = 0. \end{aligned} \tag{15}$$

Here, the variables are parameters and values of the state variables at each multiple shooting node, $s := (y, p)$, $n := \dim s$, the equalities $r_2(s) = 0$ represent the matching conditions induced by multiple shooting, $r_2(s) = (h_0^T(s), ..., h_{m-1}(s)^T)^T$. We assume that the functions $r_i : D \subset R^n \to R^{m_i}$, $i = 1, 2$, are twice-continuously differentiable.

The basic steps of the generalized Gauss-Newton algorithm with trust-region globalization applied to the nonlinear constrained least squares problem are:

1. Start with an initial guess s^0.
2. Improve the solutions iteratively by

$$s^{k+1} = s^k + \Delta s^k, \tag{16}$$

where the increment Δs^k is the solution of the linearized problem

$$\min_{\Delta s \in R^n} ||r_1(s) + J_1(s)\Delta s||_2^2, \tag{17}$$

subject to possible relaxed constraints

$$r_2(s) + J_2(s)\Delta s = (1-\alpha)r_2(s),\ 0 < \alpha \leq 1, \tag{18}$$

and a trust-region constraint

$$||\Delta s||_2^2 \leq \Delta^2. \tag{19}$$

Here, $J_i(s) = \dfrac{\partial r_i(s)}{\partial s}$, $i = 1, 2$ are the Jacobians, Δ is the trust-region radius at the $k-$th iteration and α a relaxation factor that ensures the feasibility of linear constraints and the trust region constraint in the problem (17)-(19).

Following theory of nonlinear programming, we may conclude that if the Jacobians $J_1(s)$ and $J_2(s)$ satisfy two regularity assumptions on a domain D

$$\text{rank } J_2(s) = m_2, \tag{20}$$

$$\text{rank } J = n,\ J = J(s) = \begin{pmatrix} J_1(s) \\ J_2(s) \end{pmatrix} \tag{21}$$

then a linearized trust-region problem (17)-(19) has a unique solution Δs, a unique Lagrange vector $\lambda \in R^{m_2}$, and a unique Levenberg-Marquardt parameter $\lambda_{LM} \geq 0 \in R$ satisfying the following Kuhn-Tucker conditions

$$\begin{aligned} (J_1^T(s)J_1(s) + \lambda_{LM}I)\Delta s + J_2^T(s)\lambda &= -J_1^T(s)r_1(s), \\ J_2(s)\Delta s &= -\alpha r_2(s), \end{aligned} \tag{22}$$

and the complementarity condition, namely $\lambda_{LM} = 0$ if $||\Delta s|| \leq \Delta$.

Using (22) one can easily show that under the regularity conditions (20) and (21) Δs can be formally written with the help of a solution operator $\mathcal{L}(s, \lambda_{LM}, \alpha)$:

$$\Delta s = -\mathcal{L}(s, \lambda_{LM}, \alpha)r(s), \quad r(s) = \begin{pmatrix} r_1(s) \\ r_2(s) \end{pmatrix},$$

$$\mathcal{L}(s,\lambda_{LM},\alpha)=(I \quad 0)\begin{pmatrix} J_1^T(s)J_1(s)+\mu I & J_2^T(s) \\ J_2(s) & 0 \end{pmatrix}^{-1}\begin{pmatrix} J_1^T(s) & 0 \\ 0 & \alpha I \end{pmatrix}.$$

Note that at the solution $s = s^*$ of the nonlinear problem (15) the following relations hold $\lambda_{LM} = 0$ and $\alpha = 1$ and the solution operator $\mathcal{L}(s,0,1)$ is a generalized inverse, that satisfies $\mathcal{L}(s,0,1)J\mathcal{L}(s,0,1) = \mathcal{L}(s,0,1)$. The operator $\mathcal{L}(s,0,1)$ plays a special role in statistical assessment of parameter estimation.

4.4 Evaluation of functions and Jacobians

In the course of the Gauss-Newton method the entries in the objective function and constraints and their derivatives must be evaluated frequently. The main computational effort in multiple shooting arises in the solution of the initial value problems (13) and the computation of the solution derivatives with respect to the unknowns. Efficient error controlled numerical integration methods that also deliver derivatives of the solution are required.

We use the integrator DAESOL Bauer (2001), a Backward Differentiation Formula (BDF) method with variable mesh formulas based on Newton interpolation. It uses true variable mesh error estimates for order and step size control, and a nonlinear implicit system treatment which employs strategies developed for continuation problems.

The calculation of derivatives employs "Internal Numerical Differentiation" (IND) procedures which compute derivatives of the internally generated discretization schemes. This procedure is stable in the sense of backward analysis, accurate and allows derivative error control. Moreover, it is less expensive - computing time gains of up to 80% over usual forward differences are achieved. One of the unique features responsible for the fast performance of the multiple shooting method is the adaptive accuracy strategy which keeps integration tolerances below two decimals for the most part. For a detailed discussion the reader is referred to Bock (1987).

4.5 Computing a trust-region step

To compute the trust-region step Δs^k at the point s^k we have to solve problem (17)-(19). It may happen that the linearized constraints $r_2(s^k)+J_2(s^k)\Delta s^k = 0$ and the trust-region constraint $||\Delta s^k||_2^2 \leq (\Delta^k)^2$ are inconsistent. To overcome this difficulty we relax the linear constraints and choose the relaxation factor α^k, $0 < \alpha^k \leq 1$ such that the constraints

$$\alpha r_2(s^k) + J_2(s^k)\Delta s^k = 0, \quad ||\Delta s^k||_2^2 \leq (\Delta^k)^2 \tag{23}$$

are feasible. The rules of choosing α^k will be described later.

Consider now the relaxed problem (17)-(19). Following a composite-step approach we compute the solution Δs^k of problem (17)-(19), which consists of a tangential and a normal components. This can be efficiently done by employing a block-LQ decomposition.

4.5.1 Block-LQ decomposition

The Jacobian J in the problem (17)-(19) has a very specific structure induced by the multiple shooting, which allows very effective recursive block decompositions. We describe here LQ-decomposition which is preferable for computing trust-region step because it allows to compute the trust-region step Δs^k exactly. Here, L is a lower triangular and Q is an orthogonal matrix respectively. Not only for the sake of simplicity, but rather for improving stability properties of the decomposition we handle the parameters p as constant state variables (with derivative zero) and include them in the differential variables $s_j := (y_j, p)$. The Jacobian under consideration has the form (for the sake of simplicity we omit the point s^k and the index k):

$$J = \begin{pmatrix} D_1^0 & D_1^1 & & \cdots & & D_1^m \\ G^0 & H^0 & & & & \\ & \ddots & \ddots & & 0 & \\ & 0 & & \ddots & \ddots & \\ & & & & G^{m-1} & H^{m-1} \end{pmatrix}, \quad r = \begin{pmatrix} r_1 \\ h_0 \\ \vdots \\ \vdots \\ h_{m-1} \end{pmatrix}$$

where

$$D_1^j := \partial r_1/\partial(s_j),\ G^j := \partial y(\tau_{j+1})/\partial(s_j),\ H^j := (-\mathcal{I} \quad 0), \quad j = 0, ..., m.$$

In the first step, we compute LQ-decomposition of the block $[G^0,\ H^0]$:

$$[G^0,\ H^0] = [L^0,\ 0]Q^0,$$

with Q^0 orthogonal and L^0 lower triangular and compute necessary changes in the corresponding blocks:

$$[0,\ G^1] = [T^0,\ \tilde{G}^1]Q^0,\ [D_1^0, D_1^1] = [\tilde{D}_1^0, \hat{D}_1^1]Q^0.$$

The next step of decomposition matrix is now given by

$$J^1 = \begin{pmatrix} \tilde{D}_1^0 & \hat{D}_1^1 & & \dots & & D_1^m \\ L^0 & 0 & & \dots & & 0 \\ T^0 & \tilde{G}^1 & H^1 & 0 & \dots & 0 \\ 0 & 0 & G^2 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & & 0 & G^{m-1} & H^{m-1} \end{pmatrix}$$

Now, we compute LQ-decomposition of the block $[\tilde{G}^1,\ H^1]$:

$$[\tilde{G}^1,\ H^1] = [L^1, 0]Q^1$$

and the necessary changes in the corresponding blocks:

$$[0,\ G^2] = [T^1, \tilde{G}^2]Q^1,\ [\hat{D}_1^1, D_1^2] = [\tilde{D}_1^1, \hat{D}_1^2]Q^1.$$

We proceed with this procedure until the last multiple shooting block is processed

$$[\tilde{G}^{m-1},\ H^{m-1}] = [L^{m-1},\ 0]Q^{m-1}; [\hat{D}_1^{m-1}, D_1^m] = [\tilde{D}_1^{m-1}, \tilde{D}_1^m]Q^{m-1}.$$

As a result, we get the decomposition $J = J^m Q$ with

$$J^m = \begin{pmatrix} \tilde{D}_1^0 & \tilde{D}_1^1 & & \dots & \dots & \tilde{D}_1^m \\ L^0 & 0 & & \dots & \dots & 0 \\ T^0 & L^1 & 0 & \dots & \dots & 0 \\ 0 & T^1 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & T^{m-2} & L^{m-1} & 0 \end{pmatrix}.$$

With $\Delta\tilde{s} = Q\Delta s$, the first m parts of the transformed increments (forming normal component of the trust-region step) can be computed recursively

$$\Delta\tilde{s}_0 = -(L^0)^{-1}\tilde{h}^0;$$
$$\Delta\tilde{s}_j = (L^j)^{-1}(-\tilde{h}^j - T^{j-1}\Delta\tilde{s}_{j-1}), \qquad j = 1, ..., m-1.$$

In order to find the last (tangential) part $\Delta\tilde{s}_m$ we solve the condensed problem

$$\min_{\Delta\tilde{s}_m} ||r_1 + \alpha \sum_{i=0}^{m-1} \tilde{D}_1^i \Delta\tilde{s}_i + \tilde{D}_1^m \Delta\tilde{s}_m||_2^2, \tag{24}$$

$$\text{s.t. } ||\Delta\tilde{s}_m||_2^2 \leq \bar{\Delta}^2 := \Delta^2 - \alpha^2 \sum_{i=0}^{m-1} ||\Delta\tilde{s}_i||_2^2. \tag{25}$$

This problem is solved by a classical trust-region algorithm. To recover the original increment, a recursive orthogonal transformation is performed $\Delta s = \Delta s = Q^T \Delta\tilde{s}$.

4.5.2 Reduced approach

In order to reduce the number of evaluations of derivatives to minimum, we may exploit point conditions, e.g. known initial and multipoint conditions, see Schlöder (1988). This approach is especially preferable for parameter estimation in large-scale ODE, resulting from a semidiscretization of PDEs, with only few degrees of freedom in the initial values, like in case of the problem under investigation in this paper.

Assume for simplicity that part of the equality constraints only depend on variables at one multiple shooting point. This results in entries in the linear system of the form:

$$A^i \Delta s^i = a^i, i = 0, ..., m.$$

In the first step of the reduced approach we evaluate the block A^0 and compute an LQ-decomposition

$$A^0 = [L_A^0, \ 0] Q_A^0,$$

with Q_A^0 orthogonal and L_A^0 lower triangular. Then the solution manifold can be represented as $\Delta s^0 = \Delta s_N^0 + \Delta s_T^0$,
where $\Delta s_N^0 = (Q_A^0)^T \begin{pmatrix} (L_A^0)^{-1} a^0 \\ 0 \end{pmatrix}$ and $\Delta s_T^0 = (Q_A^0)^T \begin{pmatrix} 0 \\ s_T^0 \end{pmatrix} =: \mathcal{N} s_T^0$ with s_T^0 free.

Now, we insert this solution into the first matching condition

$$G^0 \Delta s^0 + H^0 \Delta s^1 = h^0 \tag{26}$$

which then can be rewritten as

$$\mathcal{G}^0 s_T^0 - H^0 \Delta s^1 = h^0 - G^0 \Delta s_N^0, \quad \mathcal{G}^0 := G^0 \mathcal{N}. \tag{27}$$

We may apply to the matrix $[\mathcal{G}^0, H^0]$ the decomposition procedure described in the previous section, determine the solution manifold and proceed to the next multiple shooting interval.

The advantages of the reduced approach are obvious. To generate the linearized matching conditions in the form (27), only the matrix $\mathcal{G}^0$ of the directional derivatives of the initial value problem (IVP) with respect to the columns of $\mathcal{N}$ and one directional derivative $G^0 \Delta s_N^0$ have to be computed. The matrix G^0 itself is not needed. Thus, the effort for the (costly) computation of derivatives of the solution of the ODE is reduced considerably.

4.6 Computing the relaxation parameter α

The definition of $\bar{\Delta}$ (25) motivates the choice of α. If we choose

$$\alpha = \min \left(1, \frac{\sqrt{2}}{2} \frac{\Delta}{\sum_{i=0}^{m-1} ||\Delta \tilde{s}_i||_2^2} \right)$$

then $\bar{\Delta}^2 \geq \frac{1}{2}\Delta^2$, that gives us enough freedom to work on reducing the objective function.

4.7 Control of trust-region radii

The number Δ is the so-called trust-region radius that characterizes the region in which the linearized model (17)-(19) is considered to be a good approximation to the nonlinear problem. In general, the step Δs is accepted, if it produces sufficient improvement in an appropriate merit function $T(s)$. In trust-region methods, the improvement is evaluated through the ratio of the actual reduction in a merit function to the predicted reduction, that is a prediction of what the reduction in the merit function will be according to the approximation of the original problem. A traditional choice of the merit function is the so-called exact l_1-penalty function

$$T_1(s) = \frac{1}{2}||r_1(s)||_2^2 + \sum_{i=1}^{m_2} \beta_i |r_{2i}(s)| \tag{28}$$

Here, $\beta_i > 0$, $i = 1, ..., m_2$, are the penalty parameters that have to be determined in the algorithm to guarantee the global convergence of the method. Different strategies used for updating the penalty parameters and the trust-region radius and corresponding convergence theory based on classical choice of the merit function can be found e.g. in Conn et al. (2000).

However, it is well known that already in mildly ill-conditioned problems such a trust-region control strategy may be very inefficient since it may produce very small radii. Therefore we use the trust region generalization of the "restrictive monotonicity test" (RMT), see Bock et al. (2000), that has proved to be very effective in practical applications. The idea of the RMT for control of trust-region is that at s^k we consider a modified nonlinear problem:

$$\min_s \ ||r_1(s)||_2^2 + \lambda_{LM}^k ||s - s^k||_2^2, \ \ r_2(s) = (1 - \alpha^k) r_2(s^k), \tag{29}$$

for some values of λ_{LM}^k and α^k, and choose the maximal trust-region radius Δ^k for which the iterates of the simplified Gauss-Newton method, i.e. Gauss-Newton method with keeping Jacobian $J(s^k)$ fixed at all iterations, applied to (29) are contracting. This leads to the following *restrictive monotonicity test:*

Compute Δs^k as a solution of (17)-(19) with given Δ^k

$$\Delta s^k = \mathcal{L}(s^k, \alpha^k, \lambda_{LM}^k) F(s^k).$$

This corresponds to the first iteration of Gauss-Newton method applied to solve (29). The second iteration, $\tilde{\Delta} s^k$, of the simplified Gauss-Newton applied to (29) solves the linearized problem

$$\min_{\Delta s \in R^n} \ ||r_1(s^k + \Delta s^k) + J_1(s^k)\Delta s||_2^2 + \lambda_{LM}^k ||\Delta s^k + \tilde{\Delta} s^k||_2^2, \tag{30}$$
$$\text{s.t. } r_2(s^k + \Delta s^k) + J_2(s^k)\Delta s = (1 - \alpha^k) r_2(s^k),$$

and can be written as

$$\tilde{\Delta} s^k = \mathcal{L}(s^k, \alpha^k, \lambda_{LM}^k) F(s^k + \Delta s^k). \tag{31}$$

We accept the step Δs^k if

$$||\tilde{\Delta s}^k|| \leq \frac{\eta}{2} ||\Delta s^k|| \text{ for some } 0 < \eta < 2.$$

The restrictive monotonicity test has shown very good performance in practice, for the theoretical justification of the test we refer the reader to Bock et al. (2000).

4.8 Statistical sensitivity analysis for the estimates

The first results for statistical sensitivity analysis were obtained by Gauss see e.g. Gauss (1805, 1995). The discussion of the statistical sensitivity analysis for the unconstrained case can be found e.g. in Bard (1974). Here we give the results for the *constrained* least squares problems which are presented in Bock (1987) and Bock et al. (2004). If the experimental data is normally distributed then the estimated solution s^* of the parameter estimation problem is also a random variable which is normally distributed in the first order

$$s^* \sim \mathcal{N}\left(s^{\text{true}}, C\right)$$

with the (unknown) true value s^{true} as expected value and the variance-co-variance matrix C given by

$$C = C(s^*) = J^+ \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} J^{+T}. \tag{32}$$

Here, $J^+ := \mathcal{L}(s^*, 0, 1)$. The variance-covariance matrix describes the confidence ellipsoid which is an approximation of the nonlinear confidence region of the estimated variables. The matrix C can be cheaply computed using the decompositions of the Jacobians that are computed anyway in the Gauss-Newton method.

The $100\beta\%$ confidence ellipsoid $(0 \le \beta \le 1)$ can be described by

$$G_L(\beta; s^*) = \{s^* + \Delta s \mid \Delta s = -J^+(s^*) \begin{pmatrix} \eta \\ 0 \end{pmatrix}, \ \|\eta\|_2^2 \le \gamma^2(\beta)\}.$$

Here, the probability factor $\gamma(\beta)$ is given by

$$\gamma^2(\beta) = \chi^2_{n-m_1}(1-\beta)$$

where n is the dimension of s, m_1 is the dimension of the constraints of the parameter estimation problem (15), and $\chi^2_{n-m_2}(1-\beta)$ is the quantile of the

χ^2 distribution.

The diagonal elements of the covariance matrix play an important role in the statistical assessment of the estimates as well, namely they are used to compute confidence intervals $\theta_i = \sqrt{C_{ii}}\gamma(\beta)$ for each variable s_i, $i = 1, ..., n$, since

$$G_L(\beta, s^*) \subset \bigtimes_{i=1}^{n} [s_i^* - \theta_i, s_i^* + \theta_i].$$

At the solution the statistical average of the residuals, the so called common factor, can be computed by

$$\zeta = \sqrt{||r_1(s^*)||_2^2/(m_1 + m_2 - n_p)}$$

where m_1 is the number of measurements, m_2 is the number of constraints, n_p is the number of parameters. It can be used to check whether the model reproduces the measurements within the expected statistical variation.

Let us note that in multiple shooting statistical information can be computed for all variables including the values at multiple shooting nodes.

4.9 Overall algorithm

Let $\varepsilon > 0$, $\delta > 0$, $0 < \eta_1 < \eta_2 < 2$ and $0 < \gamma_1 < 1 < \gamma_2$ be specified constants. Let s^0 and Δ^0 be given.

For $k = 0, 1, 2, ...$ do until convergence (that is until $||\Delta s^k|| > \varepsilon$)

(1) Compute Δs^k, λ^k and λ_{LM}^k as the solution of problem (17)-(19).
(2) Compute $\tilde{\Delta} s^k$ as the solution of problem (30).
(3) If $||\tilde{\Delta} s^k|| > \eta_2/2||\Delta s^k||$ then do not accept the step, decrease the trust-region radius $\Delta^k := \gamma_1 \Delta^k$ and go to 1.
(4) Otherwise accept the new point $s^{k+1} = s^k + \Delta s^k$.
(5) If $||\tilde{\Delta} s^k|| > \eta_1/2||\Delta s^k||$ then increase the trust-region radius $\Delta^{k+1} = \gamma_2 \Delta^k$.

5 Application to the dollar/pound exchange rate

In this section we apply the methods described earlier to analyze the behavior of the dollar/pound exchange rate during the post-Bretton Woods period. In order to illustrate the qualitative improvements of the nonlinear model, we take the standard mean reversion model as a benchmark

$$\begin{array}{rl} \textit{Linear model} & \mu(t,X,Z) = C_0(X_1 - X) \\ \textit{Nonlinear model} & \mu(t,X,Z) = C_0(X_1 - X)(X_2 - X)(X_3 - X) \end{array}$$

The data for our investigation is taken from the International Monetary Fund's *International Financial Statistics* database, and run from March 1973 through July 2005. In particular, the dollar/pound exchange rate X is given in line "ag" (expressed as home currency per unit of foreign currency). The dynamics of the exchange rate attractors are determined by the relative behavior of the interest rates i^{US} and i^{UK} given in line "60c"

$$X_j(i^{US}, i^{UK}) = C_j \exp(i^{US})^{\alpha_{j1}} \exp(i^{UK})^{\alpha_{j2}}.$$

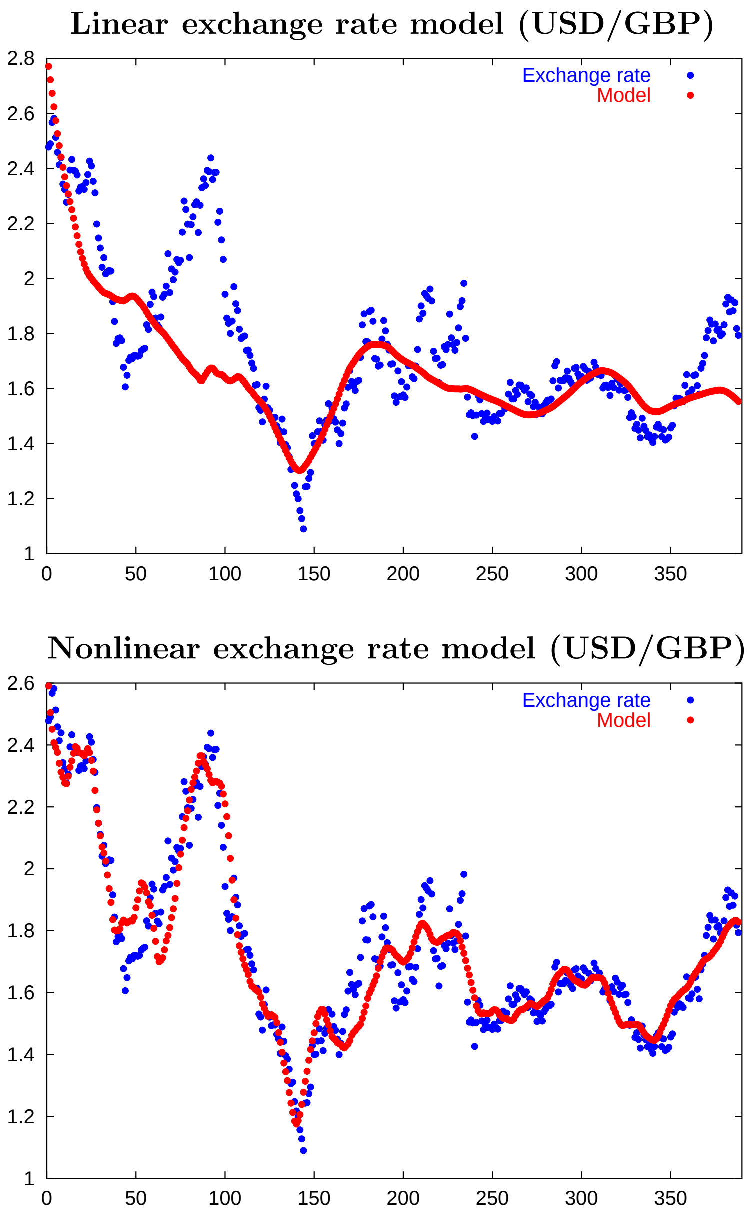
Here, we restrict ourselves to the interest rates. Adding further exchange rate fundamentals such as monetary aggregates or income measures does not improve substantially the explanatory power of the model.

Applying the presented numerical methods to the dynamics of dollar/pound exchange rates and comparing both linear and nonlinear approaches, we achieve the following results:

- By considering multiple steady states we are able to capture the historic price dynamics and distribution characteristics for the dollar/pound exchange rate. Contrary to the linear model, the generalized mean reversion process detects main turning points over a period of thirty years. Timing and direction of changes are caught surprisingly well.
- To compare the quality of the different models, we compute the root-mean-squared-error (RMSE) and the mean-average-percentage-error (MAPE) over different time periods.

  These figures demonstrate a substantial improvement of quality of pricing.
- Beside the crucial role of nonlinearities interacting with stochastic disturbances, the results highlight the importance which a stronger involvement of economic key variables has for the development of the foreign exchange rates. The set of economic data is given and not modeled. The latter is important for predictions. However, here we are mainly interested in investigating the influence of the economic variables and the effect of the nonlinearities. Therefore, the reduction to this simpler case is justified at

Fig. 4. Simulation results



The figure shows the simulation results of the dollar/pound exchange rate from 1973 to 2005 (dotted in red) and the real data (dotted in blue). In both linear (above) and nonlinear (below) approaches, the quasi-steady states depend substantially on the relative change of nominal interest rates. It can be observed that by taking into account the interplay of nonlinearity and stochastic perturbations improves the quality of pricing substantially.

Table I: Estimates of the dollar/pound exchange rate

Linear: $\mu(t, X, Z) = C_0(X_1 - X)$

Nonlinear: $\mu(t, X, Z) = C_0(X_1 - X)(X_2 - X)(X_3 - X)$

Attractors: $X_j(Z) = C_j \exp(i^{US})^{\alpha_{j1}} \exp(i^{UK})^{\alpha_{j2}}$

Linear model			**Nonlinear model**		
parameter	estimated value	± standard deviations	parameter	estimated value	± standard deviations
x_0	2.829	± 0.082	x_0	2.623	± 0.021
C_0	0.047	± 0.001	C_0	0.102	± 0.001
C_1	1.824	± 0.013	C_1	1.052	± 0.006
α_{11}	-0.034	± 0.001	α_{11}	0.025	± 0.002
α_{12}	-0.020	± 0.001	α_{12}	0.016	± 0.001
σ	0.020	± 0.001	C_2	2.180	± 0.030
			α_{21}	- 0.002	± 0.002
			α_{22}	-0.013	± 0.002
			C_3	1.722	± 0.060
			α_{31}	-0.008	± 0.004
			α_{32}	-0.027	± 0.005
			σ	0.017	± 0.001

Table II: Diagnostics of the dollar/pound exchange rate

	Linear model	**Nonlinear model**
RMSE[1]	0.101	0.056
MAPE[2]	7.85	4.49

[1]*Root Mean Squared Error,* [2]*Mean Average Percentage Error*

this state of research.

- We estimate the parameters generalizing results obtained for the quasi-steady state situation. Least squares problems constrained by partial and ordinary differential equations are already solved for real life problems such as chemical reaction systems (see e.g. Bock et al (2000)). The applied numerical method is characterized by both high accuracy and efficiency.
- In order to make forecasts, a modeling of the underlying explanatory vari-

ables is in demand. As a consequence, we get a higher dimensional system of nonlinear differential equation. In this case, the application of the presented forward Kolmogorov method depends on numerical algorithms for high dimensional problems. Recently, several different numerical methods have been developed for direct simulation for the differential equation in high dimensions, e.g. using thin grid techniques. Parameter identification in high dimension is still one of the challenges not yet overcome.

6 Conclusion

This paper treats the challenging inverse problem of the identification of the dollar/pound exchange rate mechanism; numerical results are discussed. Large price movements and multi-modal distributions can be explained by the transition between different quasi steady states which generalizes the linear mean reversion process. The attractors depend on functions of highly oscillating market fundamentals, e.g. nominal interest rates. modeling these determinants would lead to a system of stochastic differential equations and therefore to high dimensional Kolmogorov equations. This extension will be considered in future work. The methods developed for the dynamics of foreign exchange rates can be transferred to other areas in economics, e.g. the pricing of commodities. Of course, each of these fields have its own special properties. However, they also share common methodical features from a viewpoint of modeling, simulating, and validation.

References

Ait-Sahalia, Y. (1999) Transition densities for interest rate and other nonlinear diffusions, *Journal of Finance*, **4**, 1361 – 1394.

Borovkova, S., Dehling, H., Renkema, J. and Tulleken H. (2003) A potential function approach to commodity price modeling, *Computational Economics*, **22**, 139 – 161.

Bauer, I. (2001) Numerische Verfahren zur Lösung von Anfangswertaufgaben und zur Generierung von ersten und zweiten Ableitungen mit Anwendungen in Chemie und Verfahrenstechnik, PhD thesis, Universität Heidelberg.

Bard, Y. (1974) *Nonlinear parameter estimation*, Academic Press, NY.

Bilson, J.F.O. (1978) The monetary approach to the exchange rate – some empirical evidence, *IMF Staff Papers*, **25**, 48 – 75.

Bock, H.G. (1983) Recent advances in parameter identification techniques for O.D.E.. In Deuflhard and Hairer (eds), *Numerical treatment of inverse problems, progress in scientific computing*, vol. II, Birkhäuser, Boston.

Bock, H.G. (1987) Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen, *Bonner Mathematische Schriften,* **183**.

Bock, H.G., Kostina, E.A. and Schlöder, J.P. (2000) On the role of natural level functions to achieve global convergence for damped newton methods. In Powell and Scholtes (eds) *System modeling and optimization: Methods, theory and applications*, Kluwer.

Bock, H.G., Kostina, E.A. and Kostyukova, O. I. (2004) Covariance matrices for constrained parameter estimation problems, submitted to *SIAM Journal on Matrix Analysis and Applications*.

Creedy, J., Lye, J. and Martin, V.L. (1996) A nonlinear model of the real US/UK exchange rate, *Journal of Applied Econometrics*, **11**, 669 – 686.

Conn, A.R., Gould, N.I.M. and Toint, Ph.L. (2000) Trust region methods *SIAM*.

Frenkel, J.A. (1976) A monetary approach to the exchange rate: doctrinal aspects and empirical evidence, *Scandinavian Journal of Economics*, **78**, 200 – 224.

Gauss, C. F. (1995) Theory of combinations of observations least subject to errors. Original with translation, *SIAM*.

Gauss, C. F. (1805) Theoria motus corporum coelestium in sectionibus conicis solem ambientium. Original with translation, F. Perthes and J. H. Besser, Hamburg (1809).

Geman, H. (2005) *Commodities and commodity derivatives: modeling and pricing for agriculturals, metals, and energy,* John Wiley & Sons Ltd.

Kloeden, P.E. and Platen, E. (1992) *The Numerical solution of stochastic differential equations,* Springer-Verlag.

Meese, R.A. and Rogoff, K. (1983) Empirical exchange rate models of the seventies: Do they fit out of sample?, *Journal of International Economics,* **14**, 3 – 24.

MacDonald, R. and Taylor, M.P. (1993) The monetary approach to the exchange rate, *IMF Staff Papers* , **40**, 89 – 107.

Mussa, M.L. (1976) The exchange rate, the balance of payments and monetary and fiscal policy under a regime of controlled floating, *Scandinavian Journal of Economics*, **78**, 229 – 248.

Schiesser, W. (2005) *The numerical method of lines: integration of partial differential* *equations,* Academic, San Diego.

Schlöder, J.P. (1988) Numerische Methoden zur Behandlung hochdimensionaler Aufgaben der Parameteridentifizierung, *Bonner Mathematische Schriften,* **187**.

Stroock, D.W. and Varadhan, S.R.S. (1979) *Multidimensional Diffusion Processes,* Springer.

Taylor, M.P. (1995) The economics of exchange rates, *Journal of Economic Literature*, **23**, 13 – 47.