



Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 226

**A dynamic model of the financial–real interaction as a model
selection criterion for nonparametric stock market prediction**

Peter Woehrmann

January 2005

A dynamic model of the financial–real interaction as a model selection criterion for nonparametric stock market prediction

Peter Woehrmann¹

*Institute for Empirical Research in Economics, University of Zurich,
Blumlisalpstr. 10, 8006 Zurich, Switzerland.*

Abstract

Inspired by findings of low-dimensional nonlinearities and the Theorem of Takens (1983) forecasting models of financial time series are often built upon nonparametric, i.e. universal nonlinear, univariate relationships. Empirical investigations, however, are seriously contaminated by the problem of overfitting. Since statistical model selection theory in the nonlinear case is still in its infancy we would like to suggest the application of economic model selection criteria. It is a method of combining the flexibility of nonparametric regressions and important structural information in dynamic economic models. Therefore, conditions of economic models are imposed on the embedded nonlinear dynamical system to be estimated nonparametrically. In our empirical investigations we apply an univariate nonparametric forecasting model of stock returns, implemented via the Local Linear Maps of Ritter (1991), by an economic model selection criterion based on a discretized form of a continuous-time dynamic model on the interaction of real activity and asset markets. The dynamic economic model is estimated based on the Maximum Entropy inference since unobservable variables are involved. Results for monthly U.S. data show that nonparametric model selection is improved by this economic model selection criterion. On the other hand this result may be interpreted as support for the economic model.

JEL classification: C1, G1

Keywords: Economic model selection criteria, nonparametric regression, dynamic economic models, macrodynamic asset pricing

¹ I would like to thank seminar participants at the Universities of Bielefeld and Tokyo, and at the Boston College, in particular Carl Chiarella, Peter Flaschel, Thorsten Hens, Helge Ritter, and Willi Semmler for valuable comments. Financial support by the national centre of competence in research "Financial Valuation and



1 Introduction

Due to empirical findings of low-dimensional nonlinearities in financial time series that are not generated by simple parametrized functions,² in recent years empirical estimation was mainly based on econometric models that do not impose specific functional forms on the underlying stochastic processes. Among those models, both, local smoothing techniques and global approximators like neural networks³ are widely used, because they are able to represent a wide range of classes of functions.⁴ Inspired by the Theorem of Takens (1983) those models are often built upon nonparametric univariate relationships.

Empirical investigations, however, are seriously contaminated by the problem of overfitting. In particular, in the presence of short and noisy time series, this problem results in a gap between the degree of approximation and out-of-sample performance. Statistical model selection theory for nonlinear models, however, is still in its infancy.⁵

On the other hand it is well recognized that economic models are not well suited to model time series. Large macroeconomic models suffer from ad hoc type assumptions and vector autoregressions generically do not involve crucial structural relationships that might be present in small scale dynamic models. On the other hand dynamic small-scale models like the intertemporal general equilibrium models frequently omit many important variables and, therefore, could be inadequate for the purpose of approximating and forecasting time series although they capture important economic structures.

To overcome some of the deficiencies mentioned above we take an alternative approach by imposing structural economic knowledge, e.g. conditions of dynamic economic models, on the embedded nonlinear dynamical system to be estimated nonparametrically. This leads us to economic model selection criteria, i.e. choosing the nonparametric model that least violate the constraints imposed by economic models.⁶

Risk Management” is gratefully acknowledged.

² Although the application of most of those diagnostic test statistics is not free of numerical problems, the existence of nonlinear phenomena is a widely agreed result of empirical research.

³ As usual in economic and econometric literature this phrase refers to the multi-layer perceptron with backpropagation.

⁴ For most of these methods it has been shown that any functional relationship can be approximated. Note, however, that some are restricted to functions being continuously differentiable.

⁵ The nonlinear information criteria of Moody (1992) and Murata, Yoshizawa and Amari (1994) suffer from first-order approximations of Taylor-series expansions. See Woehrmann (1999a) for a Bayesian information criterion to avoid this approximation.

⁶ This, in turn, may be interpreted as support for the economic model.

A related approach to incorporate information other than the data itself into nonparametric regression is to impose constraints. An overview of this methodology is provided in Wiskott and Sejnowski (1998). A first application to financial markets can be found in Abu–Mostafa (1995) who uses symmetry hints to improve neural network forecasts of exchange rates.⁷

In our empirical investigations we implement a nonparametric univariate model of stock returns based on a local smoothing technique, the so called Local Linear Maps (LLMs) of Ritter (1991). An economic model selection criterion will be based on a dynamic model of the real activity and the asset market.⁸ Nowadays prototype dynamic models that study the interaction of real activity and asset market are often based on stochastic optimal growth models of Real Business Cycle (RBC) model type, see Kydland and Prescott (1982). Intertemporal decisions are at the heart of the RBC methodology and it is thus natural to study the asset market–output interaction in the context of those models. The asset market implications of the RBC models are, e.g., studied in Brock (1982), Danthine, Donaldson and Mehra (1992), Lettau (1997), Lettau and Uhlig (1997), Lettau, Gong and Semmler (1997).⁹ The baseline model with technology shocks as the driving force for macroeconomic fluctuations attempts to replicate basic stylized facts of the stock market such as the excess volatility of asset prices and returns and the spread between asset returns, e.g. between equity and risk–free assets. However, it turns out that the stochastic discount factor is not volatile enough to explain the high equity premium or Sharpe ratio as first pointed out in Mehra and Prescott (1985).¹⁰ Further, all standard asset pricing theories do not take into account that real activity and the stock market are mutually impacting each other. Real shocks impact asset prices but shocks to asset prices have no impact on real activity.

In this paper we, therefore, pursue an alternative macroeconomic modeling approach to model the interrelationship of the asset market and real activity. In our case the asset market will be represented by the stock market. We apply a macrodynamic model whose origin is Blanchard (1981). The Blanchard

⁷ Further related studies are those of Gallant and Tauchen (1989) who apply semi–nonparametric regressions to test an intertemporal asset pricing model and Ireland (1999) who combines vector autoregressions and the stochastic growth model.

⁸ The impact of real on financial variables is shown empirically in Fama and French (1988,1989), Fama (1990) and Schwert (1989,1990). Empirical evidence for the reverse relationship is provided in Stock and Watson (1989), Estrella and Hardouvelis (1991), Plosser and Rouwenhorst (1994) and Estrella and Mishkin (1997). Regime change models are applied in Hamilton (1989) and Hamilton and Lin (1996). Note, that those models do not capture structural economic information and are likely to suffer from specification problems.

⁹ Further intertemporal asset pricing models with production are developed in, e.g., Detemple (1986) and Cochrane (1991).

¹⁰ Time–varying asset market characteristics are spelled out in, e.g., Rouwenhorst (1995) or Woehrmann, Semmler and Lettau (1999).

variant is a perfect foresight model which exhibits saddle path stability. Only the jump to the stable branch makes the trajectories stable. Along the line of Chiarella, Semmler and Mittnik (1998) we replace the perfect foresight jump variable technique by gradual adjustments, in particular gradual expectations adjustments based on adaptive expectations. The limiting behavior of our model which admits, among others, cyclical paths generates the Blanchard model when the expectations adjust infinitely fast to yield perfect foresight as a limiting case.¹¹

We estimate the discrete-time form of the generalized Blanchard (1981) model based on the Maximum Entropy principle of Jaynes (1957) as outlined by Golan, Judge and Karp (1996). This estimation strategy is adequate since non-observable variables are involved.¹² Empirical results for U.S. output and asset market data show that the univariate nonparametric model of monthly stock returns with least mean squared prediction error admits least violation of our economic model. This fact can be interpreted as support for the economic model.

The remainder is organized as follows. In section 2 the methodology of economic model selection criteria for choosing nonparametric models and testing economic models is described. Section 3 presents the dynamic economic model and the maximum entropy approach as an adequate inference scheme. We describe its performance and evaluate it numerically. Section 4 discusses empirical results employing the economic model as a model selection criterion to selecting univariate nonparametric models of monthly U.S. stock returns. Section 5 concludes.

2 Economic model selection criteria

Consider the state of an economy represented by the bounded stochastic vector process, $z_t \in \mathbb{R}^p$, $t = 1, 2, \dots, \infty$, for some p and its law of motion,

$$g : \mathbb{R}^p \rightarrow \mathbb{R}^p, z_{t+1} = g(z_t),$$

where g is continuously differentiable or at least Lipschitz. We assume that $z_0 = \lim_{p \rightarrow \infty} z_{t_u}$ for some subsequence $t_u \rightarrow \infty$. Financial time series, $y_t \in \mathbb{R}$, $t = 1, 2, \dots, \infty$, may be interpreted as measurements of the state of the economy,

$$h : \mathbb{R}^p \rightarrow \mathbb{R}, y_t = h(z_t).$$

¹¹ Further models of this macroeconomic modeling tradition which include the financial market can be found in Flaschel, Franke and Semmler (1997).

¹² Note, however, that other techniques like, e.g., the (extended) Kalman filter could be applied.

where h is a continuously differentiable measurement function. Since it is difficult to determine all relevant components of z , the relationship between z and an embedding vector of y , $\tilde{y}_t = (y_t, y_{t-1}, \dots, y_{t-d+1})$ with embedding dimension $d > 0$ is often used to model y . To justify this approach consider

$$\phi : \mathbb{R}^p \rightarrow \mathbb{R}^d, \phi_d(z) = (h(z), h(f(z)), \dots, h(f^{d-1}(z)))$$

with $\phi_d(z) = (y_0, y_1, \dots, y_{d-1})$ for $z_0 = z$. Takens (1983) shows that ϕ_d is one-to-one and $D\phi_d(z)$ ¹³ is one-to-one at each z , if $d \geq 2p + 1$.

Therefore, econometricians assume that financial time series, y_t , are described by nonlinear functions, f , of its embedding vector, $\tilde{y}_t = (y_{t-1}, \dots, y_{t-d})$, based on the signal-plus-noise model,

$$f : \mathbb{R}^d \rightarrow \mathbb{R}, y_t = f(\tilde{y}_t) + \epsilon_t$$

with ϵ iid,¹⁴ where f is defined by conditional expectations $f(y_t) = E(y_t|\tilde{y}_t)$ with $E(\epsilon_t|\tilde{y}_t) = 0$ and $E(\epsilon f(z_t)) = 0$.

To model financial and economic variables one first has to specify a function $\psi(\tilde{y}_t, \theta)$ parameterized in θ that represents a class of functions including f .¹⁵ Then an estimation procedure has to be designed to obtain θ so as to minimize expectations of the expected loss function L , the so called risk function of Vapnik (1992),

$$R(\theta) = \int_0^\infty L(y; \psi(\tilde{y}_t, \theta)) dP(z, y)$$

with $L(y_t, \psi(\tilde{y}_t, \theta)) = \|y_t - \psi(\tilde{y}_t, \theta)\|$, joint probability $P(\tilde{y}_t, y_t)$ and $\|\cdot\|$ denoting the l_2 -norm. As $P(\tilde{y}_t, y_t)$ is not known it is suggested to minimize the empirical risk function

$$R_{\text{emp}} = T^{-1} \sum_{t=1}^T L(y_t, \psi(\tilde{y}_t, \theta))$$

based on observations y_t , $t = 1, \dots, T$.

In this work we implement nonparametric regression by the Local Linear Maps, henceforth LLMs, of Ritter (1991).¹⁶ We give only a short technical description of this method. For a lengthy note see Woehrmann (1999b). It is characterized by fast convergence and convincing generalization capabilities.¹⁷ In

¹³ Note, that D stands for the Jacobian.

¹⁴ Note, that there are also nonparametric models with heteroskedasticity.

¹⁵ We call a regression function nonparametric if it cannot be characterized by specific distributions.

¹⁶ This method is proposed independently by Stokro, Umberger and Hertz (1990) as a generalization of the widely used technique of Moody and Darken (1989).

¹⁷ Further, it offers a computational efficient way of estimating parameters and thresholds in threshold regression models simultaneously.

order to approximate function f on the basis of data y_t , $t = 1, 2, \dots, \infty$, LLMs are built up by n units $r = 1, \dots, n$, each consisting of a vector in the input space, $w_r \in \mathbb{R}^m$, a vector in the output space, $v_r \in \mathbb{R}$, and a matrix $A_r \in \mathbb{R} \times \mathbb{R}^m$. The output of an LLM for an input vector $x \in \mathbb{R}^m$ is computed as

$$\hat{y}_t = v_s + A_s(\tilde{y}_t - w_s)$$

with $s = \operatorname{argmin}_r \|\tilde{y}_t - w_r\|$. An appropriate adaptive estimation scheme for parameters A , w and v is provided by Ritter (1991),

$$\begin{aligned} \Delta w_s &= \epsilon_w(\tilde{y}_t - w_s), \\ \Delta v_s &= \epsilon_v(y_t - \hat{y}_t) + A_s \Delta w_s, \\ \Delta A_s &= \epsilon_A d_s^{-2} (y_t - \hat{y}_t)(\tilde{y}_t - w_s)' \end{aligned}$$

with $d_s = \|\tilde{y} - w_s\|$ and learning rates ϵ_w , ϵ_v and ϵ_A . Convergence of (w, v, A) to its equilibrium state (w^*, v^*, A^*) is proved in Woehrmann (1999b) using the Fokker–Planck equation approach in Ritter and Schulten (1989).

Due to the the facts that T observable variables of interest cannot be chosen very large, observations are contaminated by noise and f is possibly nonlinear and discontinuous, the task of specifying ψ^{LLM} and estimating θ turns out to be quite difficult. This is illustrated, e.g., by the bias–variance dilemma of Geman, Bienenstock and Doursat (1992).

Many theoretical investigations have been undertaken to guide the specification using statistical model selection criteria mostly in the tradition of the Akaike (1970) Information Criterion originating in information theory or the Bayesian model selection criterion of Schwarz (1978). In both cases the in-sample mean squared error (MSE) is corrected by a term depending on the complexity of f to obtain the out-of-sample MSE. For extensions that are suitable for nonlinear models see Moody (1992), Murata, Yoshizawa and Amari (1994) and Woehrmann (1999a).

Here we use the violation of an economic model to incorporate knowledge different from data itself, denoted by \mathcal{L} . We would like to interpret this also as support for the underlying economic model. Note, that tests of dynamic small-scale economic models based on traditional statistical approaches have no good asymptotic properties when the model is nonlinear and important variables are omitted.

3 The dynamic model and estimation procedure

Subsequently, we present a dynamic model on asset market and output which originates in Blanchard (1981) and is generalized in Chiarella, Semmler and

Mittnik (1998).

3.1 The dynamic model

We consider a model on asset market and output (for details see Chiarella, Semmler and Mittnik (1998)) which describes the interaction between output and stock prices in the IS–LM framework with fixed output prices¹⁸. Therefore, reaction functions for output, $y \in \mathbb{R}_+$, stock prices, $q \in \mathbb{R}_+$, and expected changes in stock returns, $x = E(\dot{q})$, are obtained as stated in the following dynamical system of differential equations.

Output adjusts to changes in aggregate expenditure, $d = aq - by + g$, where $a > 0$ and $0 \leq \beta < 1$ and $g \in \mathbb{R}_+$ is an index of fiscal expenditure. It adjusts with a delay according to

$$\dot{y} = \kappa_y (d - y) = \kappa_y (aq - by + g),$$

where $b = 1 - \beta$, $0 < b \leq 1$ and $\kappa_y > 0$ is the adjustment speed of output.

The assumption of LM equilibrium in the asset market gives $i = cy - h(m - p)$, where $c > 0$, $h > 0$ and $i \in \mathbb{R}_+$, $m \in \mathbb{R}_+$ and $p \in \mathbb{R}_+$ denote the short term interest rate, the logarithm of nominal money and the logarithm of nominal prices, respectively. Real profit is obtained as $\pi = \alpha_0 + \alpha_1 y$, where $\alpha_1 \geq 0$. Thus holding shares gives the instantaneous expected real rate of return $(x + \alpha_0 + \alpha_1 y)/q$ and the instantaneously maturing bond, i.e. the instantaneous differential between returns on shares and on short term bonds –which may allow for a (constant) equity premium– $\epsilon = (x + \alpha_0 + \alpha_1 y)/q - i$. We relax the assumption that bonds and equity are perfect substitutes, or equivalently, that any differential between them is arbitrated away instantaneously. In particular, we consider imperfect substitutability between bonds and equity by supposing that the excess demand for stocks, $q^d \in \mathbb{R}_+$, is a positive but bounded¹⁹ (sigmoid) function of the instantaneous differential ϵ , $q^d = f(\epsilon) = \tanh(\epsilon)$ with $f(0) = 0$, $f(\epsilon) \rightarrow -1$ as $\epsilon \rightarrow -\infty$ and $f(\epsilon) \rightarrow 1$ as $\epsilon \rightarrow \infty$. We suppose that the stock price adjusts to excess demand according to

$$\dot{q} = \kappa_q f(\epsilon),$$

¹⁸ Note, that this is not a severe restriction as argued in Chiarella, Semmler and Mittnik (1998).

¹⁹ This is motivated by the solution to an investor's optimization problem of intertemporal allocation of a given amount of wealth between shares and bonds so as to maximize utility arising from income, because it puts upper and lower bounds on the amounts of wealth allocated to each financial asset where borrowing is not allowed.

where $\kappa_q > 0$ is the adjustment speed of the stock market. In well functioning stock markets κ_q should be high. In the case of $\kappa_x \rightarrow \infty$, however, one obtains $(x + \alpha_0 + \alpha_1 y)/q = i$ which is an assumption in Blanchard (1981).

Expectations about the change in the stock price are assumed to be adaptive, i.e., expectations are supposed to be adjusted by

$$\dot{x} = \kappa_x(\dot{q} - x),$$

where $\kappa_x > 0$ is the adjustment speed of expectations. In the case of $\kappa_x \rightarrow \infty$ one obtains perfect foresight, i.e. $x = \dot{q}$. This is the extreme case that Blanchard (1981) considers.

In the more general case the model may be summarized in a 3-dimensional system if nonlinear differential equations,

$$\dot{y} = \kappa_y(aq - by + g), \tag{1}$$

$$\dot{q} = \kappa_q f \left(\frac{x + \alpha_0 + \alpha_1 y}{q} - cy + h(m - p) \right), \tag{2}$$

$$\dot{x} = \kappa_x \left(\kappa_q f \left(\frac{x + \alpha_0 + \alpha_1 y}{q} - cy + h(m - p) \right) - x \right). \tag{3}$$

The equilibrium of the system (1)–(3) is given by $\bar{x} = 0$ and the values (\bar{y}, \bar{q}) that solve $aq - by + g = 0$ and $(\alpha_0 + \alpha_1 y)/q = cy - h(m - p)$. We will write $\delta \equiv h(m - p)$. Two sets (\bar{y}, \bar{q}) are possible and are given by $\bar{y} = (\psi \pm \sqrt{\psi^2 - 4bc(gh(m - p) - a\alpha_0)})/(2bc)$, and $\bar{q} = (b\bar{y} - g)/a$, where $\psi \equiv gc + bh(m - p) + a\alpha_1$. Provided we assume $m > p$ there will always be at least one positive pair (\bar{y}, \bar{q}) which will be the equilibrium.

To analyze the dynamics of the system of nonlinear differential equations (1)–(3), Chiarella, Semmler and Mitnik (1998) evaluated its characteristic equation analytically and found via the usual eigenvalue criterion that κ_x may act as a bifurcation parameter with a critical value κ_x^* where the conditions of the Hopf–bifurcation theorem hold. In particular the trajectory (y_t, q_t) , $t = 1, \dots, T$ converges to a limit cycle, if $\kappa_x > \kappa_x^*$, and to a fixed point, if $\kappa_x < \kappa_x^*$.

3.2 Maximum entropy estimation

For the empirical part of the paper, i.e. in order to estimate the model based on data for output, y_t , $t = 1, 2, \dots, T$, and stock prices, q_t , $t = 1, 2, \dots, T$, we

apply the Euler scheme²⁰ to obtain (1)–(3) in discrete form,

$$\Delta y_t = \Delta t(\kappa_y(aq_{t-1} - by_{t-1} + g)) \equiv g_{y,t}, \quad (4)$$

$$\Delta q_t = \Delta t \kappa_q \bar{f} f \left(\lambda \left(\frac{x_{t-1} + \alpha_0 + \alpha_1 y_{t-1}}{q_{t-1}} - cy_{t-1} + h(m - p) \right) \right) \equiv g_{q,t}, \quad (5)$$

$$\Delta x_t = \Delta t \kappa_x \left(\frac{g_{q,t}}{\Delta t} - x_{t-1} \right) \equiv g_{x,t}. \quad (6)$$

Our estimation of the discrete system (4)–(6) is based on the Maximum Entropy Principle of Jaynes (1957) to estimate its structural parameters empirically by Golan, Judge and Karp (1996).²¹ Parameters to be estimated are summarized in

$$\begin{aligned} \varphi_y &= (\kappa_y, a, b, g), \\ \varphi_s &= (\kappa_q, \kappa_x, \alpha_0, \alpha_1, c, h, m, p) \text{ and} \\ \varphi &= (\varphi_y, \varphi_s). \end{aligned}$$

Since elements of vectors φ_y and φ_s , i.e. parameters of the output equation and the stock market equations, respectively, form two disjunct sets (4) may be estimated separately from (5) and (6). To apply the Maximum Entropy principle, support vectors,

$$Z_i = (z_{i,1}, \dots, z_{i,m})', \quad i \in \varphi,$$

and associated probability vectors,

$$P_i = (p_{i,1}, \dots, p_{i,m})', \quad i \in \varphi,$$

are defined as possible realizations of parameters and its probabilities, respectively. Hence, estimates of parameters are obtained as

$$\hat{\varphi}_i = Z_j' P_i, \quad i \in \varphi.$$

The principle of estimation is to determine $\hat{\varphi}$ with maximal uncertainty such that equilibrium conditions of the Blanchard model are satisfied. Measuring uncertainty by discrete entropy \mathcal{H} ,

$$\mathcal{H}(P_i) = -P_i' \ln P_i, \quad i \in \varphi,$$

²⁰ Although the Euler scheme only considers first order terms of the Taylor–Series expansion it performs well in comparison to improved Euler schemes, see Chiarella, Semmler and Mittnik (1998).

²¹ In contrast to standard estimation schemes the Maximum Entropy approach to estimating dynamic systems of difference equations is appropriate since it allows for incorporating variables that are not observable. However, this could also be done by, e.g., the (extended) Kalman filter.

defining the matrices $Z_j = z_{r,i}$, and $P_j = p_{r,i}$, $i \in \varphi_j$, $j \in \{y, s\}$, $r = 1, \dots, k$, and imposing regularity conditions one obtains a nonlinear program for the output equation,

$$\begin{aligned} & \max_{P_y} \mathcal{H}(P_y)' \underline{\mathbf{1}} \\ \text{s.t.} \quad & \underline{\mathbf{1}} = P_y' \underline{\mathbf{1}} \\ & \hat{\varphi}_y = Z_y' P_y \\ & \underline{\mathbf{Q}} = \Delta y_t - g_{y,t}, \quad t = 2, \dots, T, \end{aligned}$$

and a nonlinear program for the stock market equations,

$$\begin{aligned} & \max_{x, P_s} \mathcal{H}(P_s)' \underline{\mathbf{1}} \\ \text{s.t.} \quad & \underline{\mathbf{1}} = P_s' \underline{\mathbf{1}} \\ & \hat{\varphi}_s = Z_s' P_s \\ & \underline{\mathbf{Q}} = \Delta q_t - g_{q,t}, \quad t = 2, \dots, T \\ & \underline{\mathbf{Q}} = \Delta x_t - g_{x,t}, \quad t = 2, \dots, T, \end{aligned}$$

where $g_{y,t}$, $g_{q,t}$ and $g_{x,t}$ are as defined in (4)–(6). These static optimization problems are characterized by nonlinear objectives and nonlinear constraints. Therefore, we apply a projected Lagrangian algorithm based on a method due to Robinson (1972), i.e. linearly constrained subproblems are solved by augmented Lagrangians. We use the implementation in GAMS as described in Murtagh and Saunders (1982). Note, that convergence for an arbitrary starting point is not guaranteed.

Next, we test this estimation scheme by its ability to recover parameters from simulated time series with prespecified parameters regarding the output and stock market equations reported in Table 1. Employing initial conditions, $y_0 =$

Table 1
Parameters for simulations.

	φ_y	φ_s
economic structure	$a = 0.1$ $b = 0.5$	$\alpha_0 = -0.075, \bar{f} = 0.5$ $\alpha_1 = 0.15, h = 0.3$ $c = 0.1, \lambda = 3$
adjustment speed	$\kappa_y = 1$	$\kappa_q = 4, \kappa_x = 1.5$
government policy	$g = 1.3$	$\delta = 0.2$

2.1, $q_0 = 2.1$, $x_0 = 0.1$ and $T = 500$ we simulate system (4)–(6) illustrated in Fig. 1. To obtain a more realistic situation a stochastic version is simulated

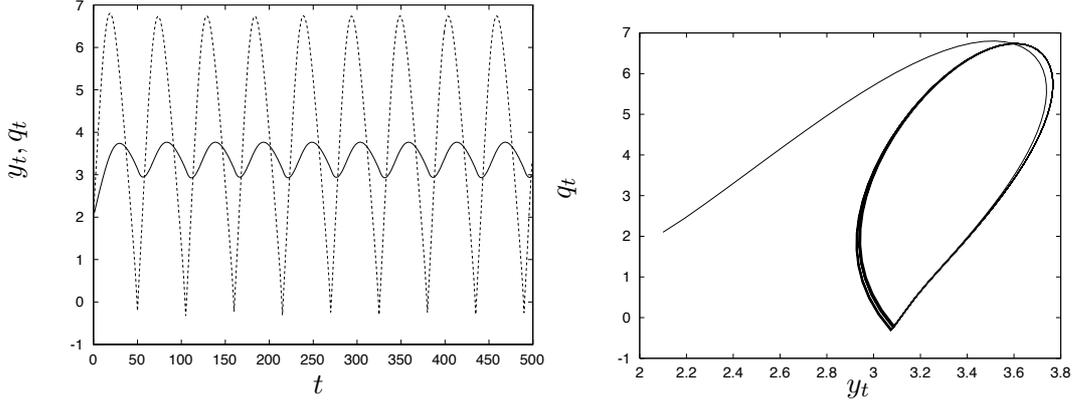


Fig. 1. Time series and phase diagrams.

by employing (7) instead of (4),

$$\Delta y_t = \Delta t(\kappa_y(aq_{t-1} - by_{t-1} + g)) + \zeta_t, \quad \zeta_t \sim N(0, \sigma_\zeta) \quad (7)$$

with $\sigma_\zeta = 0.1$. Resulting time series are illustrated in Fig. 2. Based on the

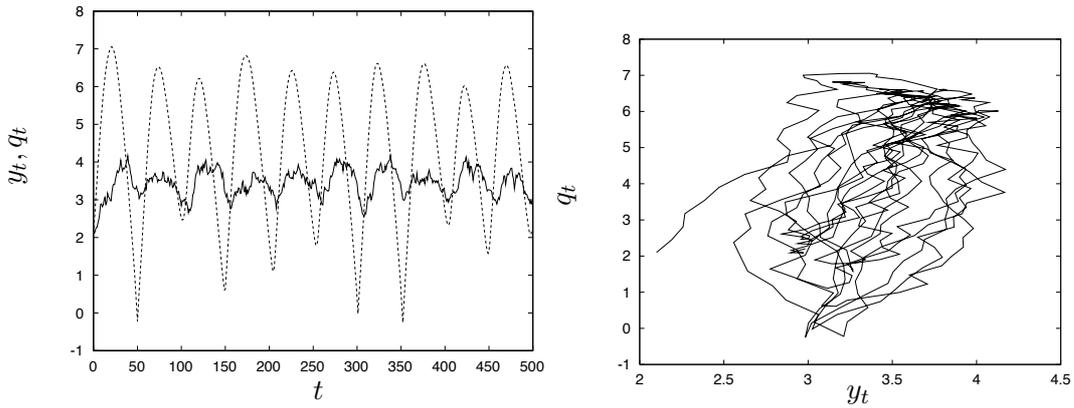


Fig. 2. Time series and phase diagrams with noise.

simulations Maximum Entropy estimation is conducted for $k = 2$. The performance is reported in Table 2. Values for parameters not reported are assumed to be known and $\hat{\cdot}$ indicates estimated parameters. Note, that the variance of estimators φ obtained from Monte Carlo simulations (not reported here) is quite low. Hence, we would like to conclude that the Maximum Entropy principle applied to our economic model works well and could be employed to real data.

Table 2
Parameter estimation for simulated time series.

	φ	$\hat{\varphi}$	$\hat{\varphi}_{\text{noise}}$	$\hat{\varphi}$	$\hat{\varphi}_{\text{noise}}$
a	0.1	0.100	0.158	0.980	0.183
b	0.5	0.500	0.575	0.490	0.665
g	1.3			1.274	1.504
κ_y	1	1.000	0.636	1.020	0.550
κ_x	4	4.000	4.000	6.030	6.596
κ_q	1.2	1.200	1.040	0.963	0.939
λ	3			1.966	1.794

4 Empirical results

Our empirical part is based on monthly U.S. data consisting of output represented by an index of industrial production and asset market represented by quotations of the S&P500 stock market index. Raw data are taken from Citibase, 1995. To obtain stationary time series we compute growth rates of industrial production and quotations of the S&P500 index denoted by y_t and q_t , respectively. We use the period from 1965.1 to 1980.12, $t = 1, 2, \dots, T$, $T = 250$, for estimation and the period from 1981.1 to 1985.12, $t = T+1, \dots, T+P$, $P = 91$, for prediction.

As motivated in section 2 we model stock returns through an univariate non-parametric model. In particular, we determine expectations of stock returns, q_t , conditioned on its embedding vector, $\tilde{q}_t = (q_{t-1}, \dots, q_{t-d})$ with $d = 10$. Therefore, we use the univariate regression model

$$\hat{q}_t^{\text{LLM}} = \psi^{\text{LLM}}(\tilde{q}_t, \theta),$$

where $\psi^{\text{LLM}}(\tilde{q}_t, \theta)$ is implemented by the Local Linear Maps of Ritter (1991).²² To evaluate in-sample and out-of-sample performance of various nonparametric models we compute

$$L_{\text{fit}} = T^{-1} \sum_{t=1}^T (q_t - \hat{q}_t^{\text{LLM}})^2 \text{ and } L_{\text{test}} = T^{-1} \sum_{t=T+1}^{T+P} (q_t - \hat{q}_t^{\text{LLM}})^2.$$

The widely recognized problem of overfitting in the case of nonparametric

²² A short description is given in section 2. For a lengthy discourse see Woehrmann (1999b).

models is investigated in Geman, Bienenstock and Doursat (1992) and is illustrated in Fig. 3. It shows the behaviour of L_{fit} and L_{test} with respect to

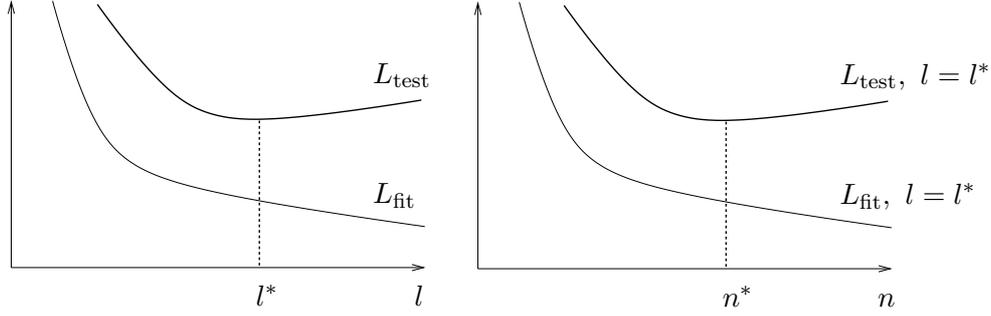


Fig. 3. Overfitting in nonparametric regression.

model complexity which is determined in the case of LLMs by $n = 1, 2, \dots$ and with respect to learning step l , $l = 1, 2, \dots$

In order to obtain an economic model selection criterion we conduct maximum entropy estimation of model (4)–(6). Estimation results are reported in Table 3. Note, that estimated values are plausible from an economic point of view.

Table 3

Parameter estimates for U.S. data.

	$\hat{\varphi}_y$	$\hat{\varphi}_s$
economic structure	$a = 0.020$ $b = 0.990$	$\alpha_0 = 0.039, \bar{f} = 0.100$ $\alpha_1 = 0.222, h = 0.505$ $c = 2.066, \lambda = 0.800$
adjustment speed	$\kappa_y = 2.885$	$\kappa_q = 1.000, \kappa_x = 7.126$
government policy	$g = 0.252$	$\delta = 0.128$

Using the stock market equation of the estimated generalized Blanchard model as discussed in section 3 we compute an economic model selection criterion,

$$\mathcal{L}_q = \sum_{t=1}^T (\Delta \hat{q}_t^{\text{LLM}} - g_{q,t})^2,$$

based on $\hat{\varphi}$ and \hat{x}_t , $t = 1, \dots, T$. This criterion which measures how much LLMs violate the estimated generalized Blanchard model will be used to evaluate different nonparametric models.

In Fig. 4 we report our estimation results for the LLMs as well as the economic model selection criteria. Boxplots summarize Monte Carlo results with 10 replications for random initial conditions (w_1, v_1, A_1) . It can be seen that choice of n with regard to \mathcal{L}_q leads to the nonparametric model with least

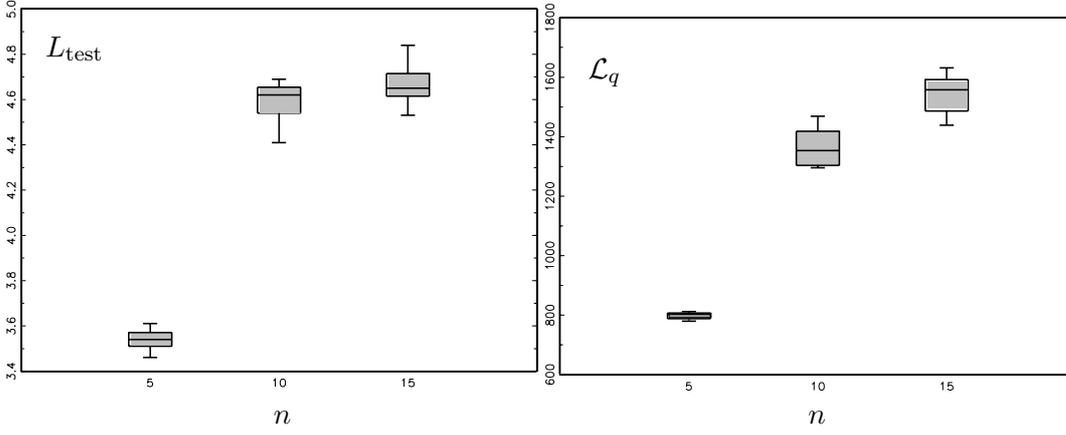


Fig. 4. Choosing complexity of LLMs by \mathcal{L}_q .

mean squared prediction error, i.e. overfitting is avoided.²³ This could also be interpreted as support for our economic model. Since we have stopped learning at $l = l^*$ we investigate whether this also could be achieved by \mathcal{L}_q .

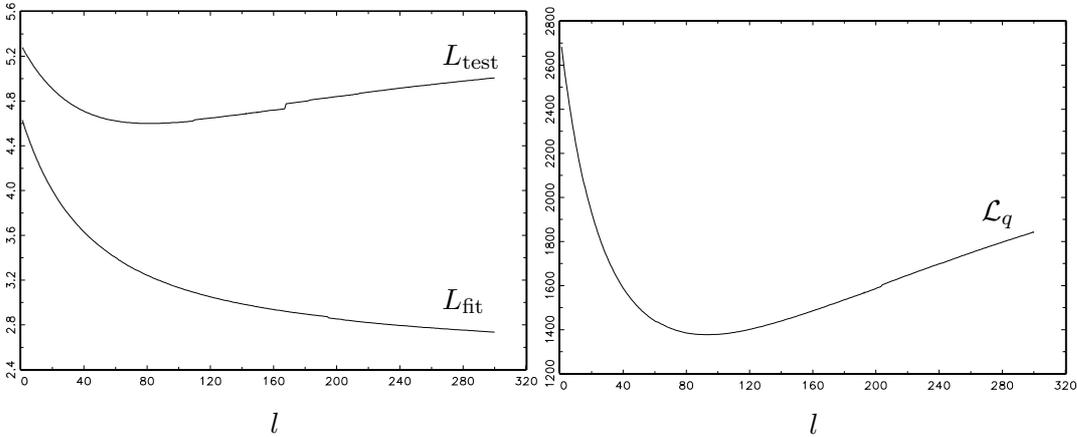


Fig. 5. Choosing no. of learning steps in LLMs ($n = 15$) by \mathcal{L}_q .

In Fig. 5 we illustrate the procedure of adaptive estimation for an LLM with $n = 15$. It provides evidence that optimal stopping may also be guided by \mathcal{L}_q .

5 Conclusion

To overcome the deficiencies of both nonparametric regressions and economic models we suggested the methodology of economic model selection criteria to combining the flexibility of the former and important structural economic knowledge of the latter approach. Therefore, we propose the method of economic model selection criteria and choose nonparametric univariate regression

²³ Note, that 71.42% of the signs of returns are predicted correctly for $n = 5$.

models that least violate conditions of dynamic economic models. This in turn can be interpreted as support for the underlying economic model.

In our empirical investigations we implement a nonparametric univariate model of stock returns based on a local smoothing technique, in particular the Local Linear Maps (LLMs) of Ritter (1991). An economic model selection criterion is based on a model of the interaction of real activity and the asset market. We here apply an alternative to RBC and rational expectations macromodels by basing our analysis on a dynamic macromodel of the interrelationship of asset market and real activity developed by Blanchard (1981) and generalized by Chiarella, Semmler and Mittnik (1998). We estimate the discrete time form of the generalized Blanchard model based on the Maximum Entropy principle of Jaynes (1957) as outlined by Golan, Judge and Karp (1996). This estimation strategy is adequate since non-observable variables are involved. Empirical results for U.S. output and asset market data show that the univariate nonparametric model of monthly stock returns with least mean squared prediction error admits least violation of our economic model. This fact could be interpreted as support for the economic model.

References

- Abu-Mostafa, Y. S., 1995, Financial applications of learning from hints, *Advances in Neural Information Processing Systems* 7, 411–418.
- Akaike, H., 1974, A new look at the statistical model selection, *IEEE Transactions on Applied Computation* AC-19, 716–723.
- Blanchard, O. J., 1981, Output, the Stock Market, and Interest Rates, *American Economic Review* 71, 132–143.
- Brock, W. A., 1982, Asset prices in a production economy, *Economics of Information and Uncertainty*, J. J. McCall, ed., University of Chicago Press, Chicago, 1–43.
- Chiarella, C., W. Semmler and S. Mittnik, 1998, *Stock Market, Interest Rate and Output: A Model and Estimations for U.S. and European Time Series Data*, Working Paper, University of Bielefeld, Germany.
- Cochrane, J., 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 209–237.
- Danthine, J., J. Donaldson and R. Mehra, 1992, The Equity Premium and the Allocation of Income Risk, *Journal of Economic Dynamics and Control* 16, 509–532.
- Detemple, J. B., 1986, Asset pricing in a production economy with incomplete information, *Journal of Finance* 41, 383–391.
- Estrella, A. and G. Hardouvelis, 1991, The Term Structure as a Predictor of Real Economic Activity, *Journal of Finance* 46, 555–576.
- Estrella, A. and F. Mishkin, 1997, The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank, *European Economic Review* 41, 1375–1401
- Fama, E., 1990, Returns and Real Activity, *Journal of Finance* 45, 1089–1107.

- Fama, E. and K. R. French, 1988, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics* 22, 3–25.
- Fama, E. and K. R. French, 1989, Business Conditions and Expected Returns on Stocks and Bonds, *Journal of Financial Economics* 25, 23–49.
- Flaschel, P., R. Franke, and W. Semmler, 1997, *Dynamic Macroeconomics: Instability, Fluctuation, and Growth in Monetary Economies*, Cambridge, MIT Press.
- Gallant, A. R. and G. Tauchen, 1989, Semiparametric estimation of conditionally constrained heterogeneous processes: Asset pricing applications, *Econometrica* 57, 1091–1120.
- Geman, S., E. Bienenstock and R. Doursat, 1992, Neural networks and the bias/variance dilemma, *Neural Computation* 4, 1–58.
- Golan, A., G. Judge and L. Karp, 1996, A Maximum Entropy Approach to Estimation and Inference in Dynamic Models, *Journal of Economic Dynamics and Control* 20, 559–582.
- Hamilton, J., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series Data and the Business Cycle, *Econometrica* 57, 357–384.
- Hamilton, J. and G. Lin, 1996, Stock Market Volatility and the Business Cycle, *Journal of Applied Econometrics* 11, 573–593.
- Ireland, P. N., 1999, A method for taking models to the data, Working Paper, Boston College and NBER, Boston, Mass.
- Jaynes, E. T., 1957, Information Theory and Statistical Mechanics, *Physics Review* 106, 620–630.
- Kydland, F. E. and E. C. Prescott, 1982, Time to build and aggregate fluctuations, *Econometrica* 50, 1345–1370.
- Lettau, M., G. Gong and W. Semmler, 1997, Statistical Estimation and Moment Evaluation of a Stochastic Growth Model with Asset Market, Working Paper, University of Bielefeld, Germany.
- Lettau, M. and H. Uhlig, 1997, Preferences, Consumption Smoothing and Risk Premia, Working Paper, Center of Economic Research, Tilburg.
- Lettau, M., 1997, Asset Prices and Business Cycles: A Diagnostic View, Working Paper, Center of Economic Research, Tilburg.
- Mehra, R. and E. C. Prescott, 1985, The Equity Premium Puzzle, *Journal of Monetary Economics* 15, 145–161.
- Moody, J. E., 1992, The effective number of parameters: An analysis of generalization and regularization in nonlinear learning systems, *Advances in Neural Information Processing Systems* 4, J. E. Moody, S. J. Hanson and R. P. Lippmann, eds., Morgan Kaufmann, San Mateo, CA, 847–854.
- Moody, J. and C. Darken, 1989, Fast learning in networks with locally tuned processing units, *Neural Computation* 1, 281–294.
- Murata, N., S. Yoshizawa and S.-I. Amari, 1994, Network information criterion – determining the number of hidden units for an artificial neural network, *IEEE Transactions on Neural Networks* 5, 865–872.
- Murtagh, B. A. and M. A. Saunders, 1982, A Projected Lagrangian Algorithm and its Implementation for Sparse Nonlinear Constraints, *Mathematical Programming Study* 16, 84–117.
- Plosser, C. and G. Rouwenhorst, 1994, International Term Structure and Real Economic Growth, *Journal of Monetary Economics* 33, 133–155.

- Ritter, H., 1991, Learning with the self-organizing map, *Artificial Neural Networks*, T. Kohonen, K. Mäkisara, O. Simula and J. Kangas, eds., Elsevier, Publishers, Amsterdam, 357–364.
- Ritter, H. and K. Schulten, 1989, Convergence properties of Kohonen’s topology conserving maps: Fluctuations, stability and dimension selection, *Biological Cybernetics* 60, 59–71.
- Robinson, S. M., 1972, A Quadratically Convergent Algorithm for General Nonlinear Programming Problems, *Mathematical Programming* 3, 145–156.
- Rouwenhorst, G. K., 1995, Asset Pricing Implications of Equilibrium Business Cycle Models, *Frontiers of business cycle research*, T. F. Cooley, ed., Princeton University Press, Princeton, 294–330.
- Schwarz, G., 1978, Estimating the dimension of a model, *Annals of Statistics* 6, 461–464.
- Schwert, G. W., 1989, Why Does Stock Market Volatility Change over Time?, *Journal of Finance* 44, 1115–1153.
- Schwert, G. W., 1990, Stock Returns and Real Activity: A Century of Evidence, *Journal of Finance* 45, 1237–1257.
- Stokro, K., D. Umberger and J. Hertz, 1990, Exploiting neurons with localized receptive fields to learn chaos, *Complex Dynamics* 4, 603–622.
- Stock, J. and Watson, 1989, New Indexes of Coincident and Leading Economic Indicators, *NBER Macroeconomic Annual*, O. J. Blanchard and S. Fischer, eds., Chicago, University of Chicago Press.
- Takens, F., 1983, Distinguishing Deterministic and Random Systems, *Nonlinear Dynamics and Turbulence*, G. I. Barenblatt, G. Iooss and D. D. Joseph, eds., Pitman Publishing, Massachusetts, 314–333.
- Vapnik, V. N., 1992, Principles of risk minimization for learning theory, *Advances in Neural Processing Systems* 4, J. E. Moody, S. J. Hanson and R. P. Lippmann, eds., Morgan Kaufmann, San Mateo, CA, pp- 831–838.
- Wiskott, L. and T. Sejnowski, 1998, Constrained optimization for neural map formation: A unifying framework for weight growth and normalization, *Neural Computation* 10, 671–716.
- Woehrmann, P., 1999a, A Bayesian information criterion for nonlinear models, Working Paper, University of Bielefeld.
- Woehrmann, P., 1999b, Nonparametric inference based on Local Linear Maps, Working Paper, University of Bielefeld.
- Woehrmann, P., W. Semmler and M. Lettau, 1999, Nonparametric estimation of time-varying asset characteristics of intertemporal asset pricing models, Working Paper, University of Bielefeld, Germany.