Studies in Nonlinear Dynamics & Econometrics

Volume 6, *Issue* 1 2002 *Article* 3

Characterizing the Degree of Stability of Non-linear Dynamic Models

Mikael Bask[*] Xavier de Luna[†]

[*]Department of Economics, Umeå University, Sweden,

[†]Department of Statistics, Umeå University, Sweden,

Characterizing the Degree of Stability of Non-linear Dynamic Models

Mikael Bask and Xavier de Luna

Abstract

The purpose of this paper is to show how the stability properties of non-linear dynamic models may be characterized and studied, where the degree of stability is defined by the effects of exogenous shocks on the evolution of the observed stochastic system. This type of stability concept is frequently of interest in economics, e.g., in real business cycle theory.

We argue that smooth Lyapunov exponents can be used to measure the degree of stability of a stochastic dynamic model. It is emphasized that the stability properties of the model should be considered when the volatility of the variable modelled is of interest. When a parametric model is fitted to observed data, an estimator of the largest smooth Lyapunov exponent is presented which is consistent and asymptotically normal. The small sample properties of this estimator are examined in a Monte Carlo study. Finally, we illustrate how the presented framework can be used to study the degree of stability and the volatility of an exchange rate.

1 Introduction

The purpose of this paper is to show how the stability properties of non-linear dynamic models may be characterized and studied, where the degree of stability is defined by the effects of exogenous shocks on the evolution of the observed stochastic system. The stability properties of models constitute a typical concern in economics. Real business cycle theory is an illustrative example (see, e.g., the review paper by Plosser, 1989), where productivity disturbances are the shocks or impulses to the economic system that are propagated across time into business cycles of various persistence.

The concept of persistence, or stability,[1] in connection with dynamic linear models is now well established and a thorough presentation is given in Chow (1975, Part I). More recently, non-linear dynamic models have attracted increased attention as they provide a richer range of dynamical behavior. The family of locally linear models has, for instance, proved popular in empirical modelling. These models are known as threshold autoregressions, where different linear autoregressive regimes explain the dependent variable (see, e.g., Granger and Teräsvirta, 1993). They have the advantage of being locally linear. Thus, the stability properties can be studied locally by direct application of the results available for linear autoregressions. Granger and Teräsvirta (1993), for example, investigate industrial production time series, where the stability properties of the different linear regimes are identified and compared.

In this paper, we present tools that make it possible to study the global stability properties of general stochastic and non-linear dynamic models. For this purpose, results available on deterministic dynamic systems are utilized. Lyapunov exponents are used to describe the dynamics of such systems (see, e.g., Eckmann and Ruelle, 1985). These exponents must be generalized in order to be applicable to stochastic models. Accordingly, we follow Nychka et al. (1992a) and consider smooth Lyapunov exponents. We give a counter-factual interpretation for these related to the persistence of an exogenous shock. This interpretation allows us to consider smooth Lyapunov exponents as a measure of the degree of stability of a stochastic dynamic model.

Another issue of interest is the relationship between the stability properties of a dynamic model and the volatility of the dependent variable modelled. The latter depends on the former, and intuitively, the more stable the model, the less volatile the dependent variable. Although this is straightforward to show for linear models, no rigorous counterpart is available, to date, when non-linear dynamics are involved. We provide, however, a heuristic argument, based on a decomposition of the volatility.

A complementary framework that can be used to study stability (which is equivalent to the Lyapunov exponent approach for linear models) is the impulse-response function. This was studied in detail in Potter (2000), who showed that there was no unique definition of such a function when non-linear dynamic models were considered. Impulse-response functions are interesting graphical tools. However, they are less appropriate when inference needs to be performed on stability measures, for instance, for model comparison.

In the next section, we start by reviewing some results on linear models in order to enhance the presentation, in Section 3, of their generalization to non-linear dynamics. Section 4 discusses how to perform inference on the smooth Lyapunov exponents. In particular, an asymptotic result in the context of parametric modelling is provided. Note that Whang and Linton (1999) showed a similar result in a non-parametric regression

[1] The term “mean reversion” is also often used in place of stability.

context. In Section 5, in order to illustrate the presented framework, we investigate the stability of the exchange rates for the Swedish Krona against the ECU and the Euro.[2] Section 6 contains a Monte Carlo study of the estimator of the smooth Lyapunov exponent. Finally, in Section 7, we conclude by discussing, in particular, how the framework presented in this paper is related to recent works where non-parametric regression techniques are used to investigate the presence of chaotic dynamics (see Dechert and Gençay, 1992, and Nychka et al., 1992a).

2 Stability of linear models

Let us begin with the simplest dynamic model,[3] an autoregression of order one for the observed time series $x_1, x_2, \ldots, x_n$,

$$x_t = \theta x_{t-1} + \varepsilon_t, \qquad t = 2, \ldots, n, \tag{1}$$

where $\{\varepsilon_t\}$ is an error term, typically independently and identically distributed with mean zero and finite variance. We assume, without loss of generality, that $E(x_t) = 0$.

Model (1) has two main components, a deterministic part that is dynamic (which we also call the *dynamics*), i.e., $x(t) = \theta x(t-1) = f(x(t-1); \theta)$,[4] and a random part (also called the *stochastics*), i.e., $\{\varepsilon_t\}$. If we focus on the dynamics, the system is explosive when $|\theta| > 1$, i.e., $|x(t)| \to \infty$ when $t \to \infty$, for any initial condition $x(0)$, except for fixed points. Furthermore, the system is conservative when $|\theta| = 1$, i.e., $x(t) = x(0)$, $\forall t$ and $\forall x(0)$. Finally, the system is stable when $|\theta| < 1$, i.e., $x(t) \to 0$ when $t \to \infty$, $\forall x(0)$.

We define the dynamic system $f(\cdot; \theta)$ as more stable, or less persistent, than the system $f(\cdot; \widetilde{\theta})$, if $|\theta| < |\widetilde{\theta}|$. Indeed, if both systems are in equilibrium and are stable, then for identical one-time shocks, the first system will converge to its equilibrium point quicker than the second system, since $x(t) = \theta^t x(0)$ and $\widetilde{x}(t) = \widetilde{\theta}^t x(0)$. Thus, when $|\theta| < 1$, $|\theta|$ is a measure of the degree of stability of the dynamic system.

The concept of stability can be generalized to higher dimensional linear autoregressions

$$x_t = \theta_1 x_{t-1} + \cdots + \theta_p x_{t-p} + \varepsilon_t, \qquad t = p+1, \ldots, n.$$

The corresponding linear dynamic system is then most conveniently described as a map from $\mathbb{R}^p$ to $\mathbb{R}^p$

$$\mathbf{f}_{\boldsymbol{\theta}} : \mathbf{x}(t-1) \to \mathbf{x}(t) = F\mathbf{x}(t-1), \tag{2}$$

where $\mathbf{x}(t-1) = [x(t-1), \ldots, x(t-p)]'$, and

$$F = \begin{pmatrix} \theta_1 & \theta_2 & \cdots & & \theta_p \\ 1 & 0 & \cdots & & 0 \\ 0 & 1 & \cdots & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & & 1 & 0 \end{pmatrix}.$$

[2] In Bask and de Luna (2001), we perform an empirical investigation of the stability of 16 European currencies using the framework developed in this paper. The political background to the launch of the Euro, and important dates in the policy decision process for the common European currency are also presented and discussed in this paper.

[3] Throughout the paper, exogenous variables that could help to explain x_t are discarded to enhance the presentation.

[4] In this paper, $\{x_t\}$ denotes both a stochastic process and a realization of a stochastic process, while $\{x(t)\}$ denotes an orbit of a deterministic dynamic system.

Although the map in (2) evolves in a p-dimensional space, it is possible to characterize its stability properties by looking at one-dimensional directions whose contraction or expansion can be described by a single parameter. The eigenvectors, $\mathbf{v}_i$, of F are defined such that

$$F\mathbf{v}_i = \delta_i \mathbf{v}_i, \qquad i = 1, \ldots, p,$$

where δ_i, the eigenvalue, is the parameter that describes the time evolution of $\mathbf{v}_i$. The eigenvectors are perpendicular by definition and form a basis of $\mathbb{R}^p$. Thus, the contraction or expansion of any direction is fully described by the modulus of the eigenvalues, $|\delta_i|$.[5]

In general, the contraction from a random state is dominated by $\delta_1 = \max_i \delta_i$. Thus, a necessary and sufficient condition for stability is that $|\delta_1| < 1$. Comparison of the stability of two p-dimensional dynamic systems can either be done by comparing the modulus of their largest eigenvalue, $|\delta_1|$, since a random shock contracts to the equilibrium point at a rate which is smaller or equal to δ_1, or by the product of the modulus of the eigenvalues, $\prod_{i=1}^{p} |\delta_i|$, as it describes the rate of contraction of a p-dimensional volume.

In the literature for dynamic systems, the quantities $\log |\delta_i|$'s are considered. These are the Lyapunov exponents for the system.

3 Non-linear models and Lyapunov exponents

Let us first consider a non-linear autoregression of order one

$$x_t = f(x_{t-1}; \boldsymbol{\theta}) + \varepsilon_t, \qquad t = 2, \ldots, n. \tag{3}$$

As previously, we define $x(t) = f(x(t-1); \boldsymbol{\theta})$. A major difference with respect to the linear case is that dissipative systems, i.e., those contracting to an attractor,[6] may be such that two orbits with arbitrary close initial conditions diverge almost always. Such systems are said to be chaotic or sensitive to initial conditions, and have been the subject of considerable research, mainly in the physical sciences (see, e.g., Eckmann and Ruelle, 1985, and Ruelle, 1989, for an introduction to chaotic dynamics). In this paper, however, we are mainly interested in non-chaotic systems, and especially in classifying them as more or less stable in a similar way as for linear systems.

If the initial conditions of two orbits are sufficiently close, it is more convenient to speak about the convergence rate of these two orbits towards each other rather than the convergence of a given orbit towards the attractor.[7] Therefore, let $x(0)$ and $\widetilde{x}(0)$ be two initial conditions that are sufficiently close. Furthermore, let $x(1) = f(x(0); \boldsymbol{\theta})$,

---

[5] Note that the autocorrelation function of the process $\{x_t\}$ is fully determined by δ_i (see, e.g., Chow, 1975, Chap. 3). The decay towards zero of these autocorrelations is therefore related to the values taken by the eigenvalues.

[6] Let A and B denote two sets such that B is the set of initial conditions $x(0)$ for which trajectories converge to A. A is then called an attractor and B its basin of attraction. A dynamic system may have several attractors, each with their own basin of attraction. Moreover, attractors may be classified into four categories (see, e.g., Ruelle, 1989, Sec. 9.4); limit points, limit cycles, quasi-periodic attractors and strange (chaotic) attractors. Stable linear systems have a unique limit point and their basin of attraction is the whole space on which the system is defined.

[7] Intuitively, if two orbits have initial conditions belonging to the same basin of attraction, and they converge to each other at a given rate, the latter also describe their rate of convergence to their common attractor.

$x(2) = f^{(2)}(x(0); \boldsymbol{\theta})$, …, $x(t) = f^{(t)}(x(0); \boldsymbol{\theta})$. Then,

$$x(t) - \widetilde{x}(t) = f^{(t)}(x(0); \boldsymbol{\theta}) - f^{(t)}(\widetilde{x}(0); \boldsymbol{\theta}) \simeq \frac{d}{dx} f^{(t)}(x(0); \boldsymbol{\theta})(x(0) - \widetilde{x}(0)),$$

where

$$\frac{d}{dx} f^{(t)}(x(0); \boldsymbol{\theta}) = \frac{d}{dx} f(x(t-1); \boldsymbol{\theta}) \frac{d}{dx} f(x(t-2); \boldsymbol{\theta}) \cdots \frac{d}{dx} f(x(0); \boldsymbol{\theta}),$$

and $|x(0) - \widetilde{x}(0)|$ is small. Thus, if $\frac{d}{dx} f(x(i); \boldsymbol{\theta})$ is of comparable size for each i, $|x(t) - \widetilde{x}(t)|$ will decrease or increase exponentially with t. The exponential rate of convergence or divergence of two orbits towards each other is then given by the *Lyapunov exponent*

$$\lambda = \lim_{t\to\infty} \frac{1}{t} \ln \left| \frac{d}{dx} f^{(t)}(x(0); \boldsymbol{\theta}) \right| = \lim_{t\to\infty} \frac{1}{t} \sum_{i=0}^{t-1} \ln \left| \frac{d}{dx} f(x(i); \boldsymbol{\theta}) \right| . \tag{4}$$

Under suitable technical conditions (Ruelle, 1989, Chap. 9), this limit exists and is independent of $x(0)$ almost surely (with respect to the invariant measure induced by f). Although we omit it in the notation, λ is a function of the parameter $\boldsymbol{\theta}$. Note that we also can write

$$|x(t) - \widetilde{x}(t)| \simeq \exp(\lambda t) \, |x(0) - \widetilde{x}(0)| \, .$$

The Lyapunov exponent generalizes the slope of the linear autoregression, θ in (1), for which $\lambda = \ln |\theta|$. We have a conservative system when $\lambda = 0$, and when $\lambda < 0$ a bounded system is said to be stable. When $\lambda > 0$, the system is said to be chaotic if it is bounded, otherwise it is explosive. Chaotic and stable systems are dissipative since any orbit converges to the attractor, although the former have a random like nature and are not predictable in practice.

The stability of two different non-linear dynamic models may be directly compared by looking at their respective Lyapunov exponents. A system with a large and negative Lyapunov exponent will be classified as more stable than a system with a negative Lyapunov exponent that is close to zero.

The Lyapunov exponent in (4) is defined with respect to the measure induced by f, and not the one induced by the distribution of the stochastic process, $\{x_t\}$. Therefore, this definition of the Lyapunov exponent does not take into account the influence of the stochastics. For linear systems, however, this issue does not arise since $\frac{d}{dx} f(x(i); \theta) = \theta$ for every i, i.e., the derivative is constant at any state of the process.

For non-linear stochastic systems, define

$$\overline{\lambda} = \lim_{t\to\infty} \frac{1}{t} \ln \left| \frac{d}{dx} f^{(t)}(x_0; \boldsymbol{\theta}) \right| = \lim_{t\to\infty} \frac{1}{t} \sum_{i=0}^{t-1} \ln \left| \frac{d}{dx} f(x_i; \boldsymbol{\theta}) \right| , \tag{5}$$

where we have used the convention $x_0 = x(0)$ and $x_1 = f(x_0) + \varepsilon_1$, and so on.[8] Conditions for the existence and almost sure independence of x_0 can be found in McCaffrey et al. (1992). Although we omit it again in the notation, $\overline{\lambda}$ is a function of the parameter $\boldsymbol{\theta}$. The Lyapunov exponent in (5) is a smoothed version of the exponent in (4), and while the latter may be a non-continuous function of $\boldsymbol{\theta}$ (Ruelle, 1989, Sec. 9.5), the former is continuous (Nychka et al., 1992b). Therefore, we call $\overline{\lambda}$ a *smooth Lyapunov exponent*.

[8] Note that $\overline{\lambda}$ differs essentially from λ in being evaluated at $(x_1, x_2, \ldots)$, an observed realization of a stochastic process, and not at $(x(1), x(2), \ldots)$, an orbit of the dynamic part of the model.

There is no real consensus in the literature concerning which definition of the Lyapunov exponent should be considered as appropriate for stochastic processes (see Tong, 1995, and the discussion therein for some divergent views on this issue). Our standpoint is that the definitions discussed in the literature are complementary and their appropriateness depends on the purpose of the analysis. For instance, when models are fitted for predictive purposes then the measures defined in Yao and Tong (1994) and Wolff (1992) are relevant, whereas when models are intended to be used for theoretic reasoning, the smooth Lyapunov exponent may be more useful, as is clarified in the next two sections.

3.1 Counter-factual interpretation of the smooth Lyapunov exponent

Let $x_0 = x(0) + \eta$ and $\widetilde{x}_0 = x(0)$ be two initial conditions, where $\eta > 0$ is a one-time shock to the dynamic system. Assume that we observe a realization of the process $\{x_t\}$, generated by model (3), with the initial conditions x_0 and $x_1 = f(x_0) + \varepsilon_1$. A set of realized values for $\varepsilon_1, \ldots, \varepsilon_t$ corresponds to this observed series. A question of interest is then the persistence of the shock η in the system. To study this issue, we define the non-observed series $\{\widetilde{x}_t\}$ with the initial condition $\widetilde{x}_0$, where the same set of realized values for $\varepsilon_1, \ldots, \varepsilon_t$, as when generating $\{x_t\}$ are used. Basing the non-observed series on the same noise realizations as $\{x_t\}$ is justified by the fact that the process $\{\varepsilon_t\}$ is exogenous and independent of the initial condition. In other words, the persistence of η is studied keeping all other disturbances fixed.

The asymptotic persistence of $\eta > 0$ is obtained by deriving the convergence rate of $x_t - \widetilde{x}_t$. Writing $f_{\boldsymbol{\theta},\varepsilon}(x) = f(x; \boldsymbol{\theta}) + \varepsilon$, where ε is now a parameter, and

$$g^{(t)}(x; \boldsymbol{\theta}, \boldsymbol{\varepsilon}_t) = \underbrace{f_{\boldsymbol{\theta},\varepsilon_t} \circ f_{\boldsymbol{\theta},\varepsilon_{t-1}} \circ \cdots \circ f_{\boldsymbol{\theta},\varepsilon_1}(x)}_{t \text{ times}},$$

where $\boldsymbol{\varepsilon}_t = (\varepsilon_1, \ldots, \varepsilon_t)$, we have, for $|\eta|$ small,

$$\begin{aligned} x_t - \widetilde{x}_t &= g^{(t)}(x_0; \boldsymbol{\theta}, \boldsymbol{\varepsilon}_t) - g^{(t)}(\widetilde{x}_0; \boldsymbol{\theta}, \boldsymbol{\varepsilon}_t) \\ &\simeq \frac{d}{dx} g^{(t)}(x_0; \boldsymbol{\theta}, \boldsymbol{\varepsilon}_t)(x_0 - \widetilde{x}_0), \end{aligned}$$

where

$$\frac{d}{dx} g^{(t)}(x_0; \boldsymbol{\theta}, \boldsymbol{\varepsilon}_t) = \frac{d}{dx} f(x_{t-1}; \boldsymbol{\theta}) \frac{d}{dx} f(x_{t-2}; \boldsymbol{\theta}) \cdots \frac{d}{dx} f(x_0; \boldsymbol{\theta}).$$

Then,

$$|x_t - \widetilde{x}_t| \simeq \exp(\overline{\lambda} t) \, |x_0 - \widetilde{x}_0| \, .$$

Thus, we can interpret $\overline{\lambda}$ as the exponential rate of convergence (or divergence) of the series $\{x_t\}$ towards (or from) the non-observed series $\{\widetilde{x}_t\}$.

This counter-factual interpretation of $\overline{\lambda}$ can be translated as follows: a one-time shock has shorter term consequences for a stochastic dynamic system with a large and negative $\overline{\lambda}$ than for a system with a negative $\overline{\lambda}$ that is close to zero. The latter system is, in this sense, less stable than the former.[9]

Note that λ characterizes another, noise free, dynamic system than the one of interest. In this respect it has an interesting although different interpretation than $\overline{\lambda}$. Unless the variability of the stochastic term $\{\varepsilon_t\}$ is negligible when compared to the variability induced by the dynamics, $\overline{\lambda}$ and λ can be very different.

[9] A related argument in favor of $\overline{\lambda}$ was given by Herzel, Ebeling and Schulmeister (1987).

3.2 Decomposition of the marginal volatility

Ideally it would be desirable to have a decomposition of the volatility, $Var(x_t)$, showing explicitly its dependence on the exogenous volatility, $Var(\varepsilon_t)$, and the stability properties discussed above. For model (1) we can write, assuming $|\theta| < 1$,

$$Var(x_t) = \frac{Var(\varepsilon_t)}{1-\theta^2} = \frac{Var(\varepsilon_t)}{1-\exp(\overline{\lambda})^2}.$$

Thus, the marginal volatility increases both when the exogenous shocks $\{\varepsilon_t\}$ become more volatile and when the system becomes less stable, i.e., $\overline{\lambda}$ approaches zero from below. Such an explicit decomposition is not available in the general non-linear case. However, considering again an infinitesimal shock η, say with mean zero and variance $\sigma_\eta^2 \to 0$, at time $t-1$ such that $\widetilde{x}_t = f(x_{t-1}+\eta) + \varepsilon_t$. Then, under regularity conditions,

$$\widetilde{x}_t \simeq f(x_{t-1}) + \frac{df(x_{t-1})}{dx}\eta + \varepsilon_t,$$

and

$$Var(\widetilde{x}_t|x_{t-1}) \simeq \left[\frac{df(x_{t-1})}{dx}\right]^2 \sigma_\eta^2 + \sigma_\varepsilon^2.$$

This is a simplified version of the decomposition theorem given in Yao and Tong (1994), where the mean squared prediction error was considered. The latter expression describes what the effect of a shock η at time $t-1$ would be on the conditional variance. This effect (an increase of volatility by $\left[\frac{df(x_{t-1})}{dx}\right]^2 \sigma_\eta^2$) is local since it depends, in general, on x_{t-1}. An average effect is then $E\left\{\left[\frac{df(x_{t-1})}{dx}\right]^2\right\}\sigma_\eta^2$. The stability measure $\overline{\lambda}$ is a lower bound for $\frac{1}{2}\ln E\left\{\left[\frac{df(x_{t-1})}{dx}\right]^2\right\}$ (by Jensen's inequality), indicating that if a stochastic system becomes less stable, i.e., $\overline{\lambda}$ increases, then its volatility will probably increase. This, however, cannot be guaranteed in contrast with linear models.

3.3 Higher order models

The two different definitions of the Lyapunov exponents can be generalized to higher dimensional non-linear autoregressions

$$x_t = f(x_{t-1}, \cdots, x_{t-p}; \boldsymbol{\theta}) + \varepsilon_t, \qquad t = p+1, \ldots, n, \tag{6}$$

where the corresponding dynamics is most conveniently described as a map from $\mathbb{R}^p$ to $\mathbb{R}^p$

$$\mathbf{f}_{\boldsymbol{\theta}} : \mathbf{x}(t-1) \to \mathbf{x}(t) = F(\mathbf{x}(t-1); \boldsymbol{\theta}),$$

where $\mathbf{x}(t-1) = [x(t-1), \ldots, x(t-p)]'$.

Denote by $Df(\mathbf{x};\boldsymbol{\theta})$ the Jacobian matrix that contains the partial derivatives of F evaluated at $\mathbf{x}$, and consider the linearized system

$$\mathbf{x}(t) - \widetilde{\mathbf{x}}(t) \simeq Df^{(t)}(\mathbf{x}(0);\boldsymbol{\theta})(\mathbf{x}(0) - \widetilde{\mathbf{x}}(0)),$$

where $Df^{(t)}(\mathbf{x}(0);\boldsymbol{\theta}) = Df(\mathbf{x}(t-1);\boldsymbol{\theta})Df(\mathbf{x}(t-2);\boldsymbol{\theta})\cdots Df(\mathbf{x}(0);\boldsymbol{\theta})$. We can now use the arguments used for a p-dimensional linear system and look at the eigenvalues $\delta_i(t, x(0);\boldsymbol{\theta})$

of $Df^{(t)}(\mathbf{x}(0))$, where $|\delta_1(t,x(0);\boldsymbol{\theta})| \geq |\delta_2(t,x(0);\boldsymbol{\theta})| \geq \cdots \geq |\delta_p(t,x(0);\boldsymbol{\theta})|$. The Lyapunov spectrum can then be defined as

$$\lambda_i = \lim_{t\to\infty} \frac{1}{t} \ln |\delta_i(t,x(0);\boldsymbol{\theta})|, \qquad i = 1,\dots,p.$$

Under suitable technical conditions (Ruelle, 1989, Chap. 9), these limits exist and are independent of $x(0)$ almost surely (with respect to the measure induced by f). As before, we can define the smooth Lyapunov spectrum as

$$\overline{\lambda}_i = \lim_{t\to\infty} \frac{1}{t} \ln |\delta_i(t,\mathbf{x}_0;\boldsymbol{\theta})|, \qquad i = 1,\dots,p,$$

where $\delta_i(t,\mathbf{x}_0;\boldsymbol{\theta})$ are the eigenvalues of $Df(\mathbf{x}_{t-1};\boldsymbol{\theta})Df(\mathbf{x}_{t-2};\boldsymbol{\theta})\cdots Df(\mathbf{x}_0;\boldsymbol{\theta})$. Again, for its existence and independence of $\mathbf{x}_0$, see McCaffrey et al. (1992).

4 Inference on smooth Lyapunov exponents

We here assume that, based on the observed time series $x_1,\dots,x_n$, a consistent estimator $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is available. Then, consider the estimator of the Lyapunov exponent for the order one case

$$\widehat{\overline{\lambda}} = \frac{1}{m} \sum_{i=1}^{m} \ln \left| \frac{d}{dx} f(x_i;\widehat{\boldsymbol{\theta}}) \right|, \tag{7}$$

and for Lyapunov spectrum for the order $p > 1$ case

$$\widehat{\overline{\lambda}}_i = \frac{1}{m} \ln |\delta_i(m,\mathbf{x}_1;\widehat{\boldsymbol{\theta}})|, \qquad i = 1,\dots,p, \tag{8}$$

where in both cases $\mathbf{x}_1,\dots,\mathbf{x}_m$ is a sub-sample of size m from the original sample. The value of m must be chosen in accordance with assumption (A5) made below. Moreover, $\delta_i(m,\mathbf{x}_1;\widehat{\boldsymbol{\theta}})$ are the eigenvalues of $Df(\mathbf{x}_m;\widehat{\boldsymbol{\theta}})Df(\mathbf{x}_{m-1};\widehat{\boldsymbol{\theta}})\cdots Df(\mathbf{x}_1;\widehat{\boldsymbol{\theta}})$.

We assume that the estimator $\widehat{\boldsymbol{\theta}}$ has the following common and desirable asymptotic properties:

(A1) $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \overset{n\to\infty}{\sim} N(0,\Sigma)$, where Σ is the $p \times p$ variance-covariance matrix. In particular, $\widehat{\boldsymbol{\theta}}$ is $\sqrt{n}$-consistent for $\boldsymbol{\theta}_0$, the unknown and "true" value of $\boldsymbol{\theta}$: $\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 = O_p(n^{-1/2})$.

Furthermore,

(A2) Denote by $\nabla_{\boldsymbol{\theta}} \frac{d}{dx} f(x;\boldsymbol{\theta})$ the gradient vector of partial derivatives $\frac{d}{dx} f(x;\boldsymbol{\theta})$ with respect to the elements of $\boldsymbol{\theta}$. Then $\sup_x |\nabla_{\boldsymbol{\theta}} \frac{d}{dx} f(x;\boldsymbol{\theta}_0)|$ exists, is bounded, and is also continuous in $\boldsymbol{\theta}_0$.

(A3) The stochastic process $\{x_t\}$ is strongly mixing with a mixing number of size $r/(r-1)$ for some $r > 1$ (see, e.g., White and Domowitz, 1984, for the definition of strongly mixing).

(A4) $E(|\ln|\frac{d}{dx} f(x_t;\boldsymbol{\theta}_0)||^{2r}) < \infty$ is bounded uniformly in i, and $\lim_{n\to\infty} \overline{\sigma}^2_{s,n} = \overline{\sigma}^2 < \infty$ uniformly in s, where $\overline{\sigma}^2_{s,n} = Var(n^{-1/2} \sum_{i=s+1}^{s+n} \ln |\frac{d}{dx} f(x_i;\boldsymbol{\theta}_0)|)$.

(A5) $\max_{i=1,\dots,n}(|\frac{d}{dx}f(x_i;\boldsymbol{\theta}_0)|^{-1}) = (\min_{i=1,\dots,n}|\frac{d}{dx}f(x_i;\boldsymbol{\theta}_0)|)^{-1} = O_p(n^{\delta})$ for some $\delta \geq 0$.

Sufficient conditions for assumption (A1) to hold may be found, for example, in White and Domowitz (1984). Assumptions (A2) and (A4) are regularity conditions, and assumption (A5) allows $\frac{d}{dx}f(x;\boldsymbol{\theta}_0)$ to be zero for some x by choosing $\delta > 0$. This is discussed further in Whang and Linton (1999). Finally, sufficient conditions for assumption (A3) to hold for model (6) are provided by Chan and Tong (1994, Theorem 1).

Proposition 1 *Assuming (A1)-(A5) and $m \to \infty$ at most with pace $m = o(n^{1/(1+2\delta)})$, we have asymptotically ($n \to \infty$) that*

$$\sqrt{m}(\widehat{\overline{\lambda}} - \overline{\lambda}) \sim N(0, V),$$

where $V = \lim_{n\to\infty} m^{-1} Var(\sum_{i=1}^m \ln|\frac{d}{dx}f(x_i;\boldsymbol{\theta}_0)|)$.

Proof. Write

$$\begin{aligned}
\sqrt{m}(\widehat{\overline{\lambda}} - \overline{\lambda}) &= \frac{1}{m^{1/2}}\sum_{i=1}^m \left[\ln\left(\frac{d}{dx}f(x_i;\widehat{\boldsymbol{\theta}})\right) - \ln\left(\frac{d}{dx}f(x_i;\boldsymbol{\theta}_0)\right)\right] + \\
&\quad \frac{1}{m^{1/2}}\sum_{i=1}^m \left[\ln\left(\frac{d}{dx}f(x_i;\boldsymbol{\theta}_0)\right) - \overline{\lambda}\right] \\
&= A_m + B_m.
\end{aligned}$$

We first show that $A_m = o_p(1)$. By a Taylor expansion about $\frac{d}{dx}f(x_i;\boldsymbol{\theta}_0)$, we have that

$$|A_m| = \left|\frac{1}{m^{1/2}}\sum_{i=1}^m \frac{1}{\frac{d}{dx}f(x_i;\widetilde{\boldsymbol{\theta}})}\left[\frac{d}{dx}f(x_i;\widehat{\boldsymbol{\theta}}) - \frac{d}{dx}f(x_i;\boldsymbol{\theta}_0)\right]\right|,$$

where $\widetilde{\boldsymbol{\theta}}$ lies between $\boldsymbol{\theta}_0$ and $\widehat{\boldsymbol{\theta}}$. Furthermore,

$$\begin{aligned}
|A_m| &\leq \frac{1}{m^{1/2}}\sup_x \left|\frac{d}{dx}f(x_i;\widehat{\boldsymbol{\theta}}) - \frac{d}{dx}f(x_i;\boldsymbol{\theta}_0)\right| \sum_{i=1}^m \frac{1}{\left|\frac{d}{dx}f(x_i;\widetilde{\boldsymbol{\theta}})\right|} \\
&\leq \frac{m^{1/2\delta}}{n^{1/2}} n^{1/2}\sup_x \left|\frac{d}{dx}f(x;\widehat{\boldsymbol{\theta}}) - \frac{d}{dx}f(x;\boldsymbol{\theta}_0)\right| \frac{1}{m^{\delta}\min_i\left|\frac{d}{dx}f(x_i;\widetilde{\boldsymbol{\theta}})\right|}.
\end{aligned}$$

By a Taylor expansion, we have that $\frac{d}{dx}f(x;\widehat{\boldsymbol{\theta}}) - \frac{d}{dx}f(x;\boldsymbol{\theta}_0) = \nabla_{\boldsymbol{\theta}}\frac{d}{dx}f(x_i;\widetilde{\boldsymbol{\theta}})'(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$, where $\widetilde{\boldsymbol{\theta}}$ lies between $\boldsymbol{\theta}_0$ and $\widehat{\boldsymbol{\theta}}$. Therefore,

$$n^{1/2}\sup_x\left|\frac{d}{dx}f(x;\widehat{\boldsymbol{\theta}}) - \frac{d}{dx}f(x;\boldsymbol{\theta}_0)\right| = n^{1/2}\sup_x\left|\nabla_{\boldsymbol{\theta}}\frac{d}{dx}f(x_i;\widetilde{\boldsymbol{\theta}})'(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)\right| = O_p(1),$$

by assumptions (A1) and (A2). Moreover, by assumption (A5), $\left[m^{\delta}\min_i\left|\frac{d}{dx}f(x_i;\widetilde{\boldsymbol{\theta}})\right|\right]^{-1} = O_p(1)$ (see Whang and Linton, 1999, Appendix A). Hence, $A_m = o_p(1)$ since $n^{-1/2}m^{1/2+\delta} = o(1)$ is assumed.

Turning to B_m, by assumptions (A3) (according to White and Domowitz, 1984, Lemma 2.1, $\ln\left|\frac{d}{dx}f(x_i;\boldsymbol{\theta}_0)\right|$ is strongly mixing) and (A4), we have the central limit theorem (CLT) result (see White and Domowitz, 1984, Theorem 2.4):

$$B_m \overset{n\to\infty}{\sim} N(0,V),$$

which concludes the proof. ■

Remark 1 *We consider in Proposition 1 sub-samples of size $m<n$ in order to be able to relax the restrictive assumption $\delta=0$. On the other hand, even if $\delta=0$, m must tend to infinity at a slower rate than n ($m=o(n)$) for A_m to vanish in the above proof. However, assuming that $\delta=0$, as well as further regularity conditions (allowing a first order Taylor expansion of A_n about $\boldsymbol{\theta}_0$), one may also show the asymptotic normality of $\widehat{\widehat{\lambda}}$ when $m=n$. We do not present the additional necessary assumptions or the proof of this result here,*[10] *but we give the asymptotic variance for the case $m=n$ (this variance takes into account the variability of $\widehat{\boldsymbol{\theta}}$):*

$$W = V + \lim_{n\to\infty} E(\xi'\Sigma\xi), \tag{9}$$

where

$$\xi = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\nabla_{\boldsymbol{\theta}}\frac{d}{dx}f(x;\boldsymbol{\theta}_0)}{\frac{d}{dx}f(x;\boldsymbol{\theta}_0)}.$$

Remark 2 *Whang and Linton (1999) gave the asymptotic distribution of $\widehat{\widehat{\lambda}}$ and $\widehat{\widehat{\lambda}}_1$ for a non-parametric estimation of the dynamics. Proposition 1 above can also be generalized to the case $p>1$ for $\widehat{\widehat{\lambda}}_1$, although both the regularity conditions and the technical developments are more involved.*

Remark 3 *Because the mixing conditions in assumption (A3) are necessary for the estimation of $\boldsymbol{\theta}$, we also find it natural to use them to obtain the CLT for B_m in the above proof. However, it is also possible to use the CLT obtained by Bailey et al. (1997). It is based on more explicit assumptions about the dynamics and stochastics of the process. For instance, they argue that the CLT depends on the non-linearity of the dynamics and on the presence of the stochastics component.*

A consistent estimator of V can be obtained by a so-called block estimator (Bartlett, 1950), where a block size b is chosen to compute

$$l_i = \frac{1}{b}\sum_{j=i}^{i+b-1}\ln\left|\frac{d}{dx}f(x_i;\widehat{\boldsymbol{\theta}})\right|, \qquad i=1,\ldots,n-b+1,$$

with which an estimator of V is obtained as

$$\widehat{V} = \frac{1}{n-b+1}\sum_{i=1}^{n-b+1}(l_i-\bar{l})^2,$$

where $\bar{l} = \frac{1}{n-b+1}\sum_{i=1}^{n-b+1} l_i$ or $\bar{l}=\widehat{\widehat{\lambda}}$. Finally, the second element of W in (9) can be estimated by $\widehat{\xi}'\Sigma\widehat{\xi}$, where $\widehat{\xi}$ is ξ evaluated at $\widehat{\boldsymbol{\theta}}$.

[10] In fact, we believe that the assumption $\delta=0$ should be possible to relax even when using the full sample. The Monte Carlo study presented in Section 6 below gives evidence for this.

5 Application: Volatility and stability of exchange rates

In order to illustrate the framework presented, we consider the exchange rates for the Swedish Krona against the common European currency. Two time periods are considered: Period I ranges from May 2, 1998 (i.e., the date when the Brussels summit decided the member states of the EMU) to December 31, 1998, and Period II ranges from January 1, 1999 (i.e., the EMU is implemented) to March 31, 2001.[11] For a given a time period, a parametric model (to be more specific, a polynomial autoregression on a projected space, see de Luna, 1998) is identified and fitted

$$x_t = \alpha_0 + \sum_{i=1}^{d} \alpha_i(\theta_1 x_{t-1} + \cdots + \theta_p x_{t-p})^i + \varepsilon_t. \tag{10}$$

Smooth Lyapunov exponents are then estimated from the fitted model.[12]

The idea here is that if a structural change has occurred when the Euro was launched, then the models obtained for Periods I and II may display different properties. Models ($1 \leq d \leq 3$ and $0 \leq p \leq 7$) were identified using the AIC criterion (Akaike, 1974), and fitted consistently with non-linear least squares. The results are as follows:

Period I (242 observations):

$\widehat{d} = 3 : \widehat{\alpha}_0 = 9.5 \times 10^{-4}$, $\widehat{\alpha}_1 = -0.06$, $\widehat{\alpha}_2 = -5.68$, $\widehat{\alpha}_3 = 307.9$

$\widehat{p} = 5 : \widehat{\theta}_1 = 1.1286$, $\widehat{\theta}_2 = -0.250$, $\widehat{\theta}_3 = 0.406$, $\widehat{\theta}_4 = -0.773$, $\widehat{\theta}_5 = -1.10$

$\widehat{Var}(\varepsilon_t) = 21.8 \times 10^{-6}$

Explained variance: $\left[\widehat{Var}(x_t) - \widehat{Var}(\varepsilon_t)\right] / \widehat{Var}(x_t) = 4\%$

The maximum smooth Lyapunov exponent: $\widehat{\overline{\lambda}}_1 = -0.44\ (-0.66, -0.33)$

Period II (820 observations):

$\widehat{d} = 2 : \widehat{\alpha}_0 = -1.1 \times 10^{-4}$, $\widehat{\alpha}_1 = 0.09$, $\widehat{\alpha}_2 = 3.90$

$\widehat{p} = 6 : \widehat{\theta}_1 = 0.40$, $\widehat{\theta}_2 = -1.24$, $\widehat{\theta}_3 = 0.22$, $\widehat{\theta}_4 = -0.10$, $\widehat{\theta}_5 = 0.44$, $\widehat{\theta}_6 = -0.95$

$\widehat{Var}(\varepsilon_t) = 9.5 \times 10^{-6}$

Explained variance: $\left[\widehat{Var}(x_t) - \widehat{Var}(\varepsilon_t)\right] / \widehat{Var}(x_t) = 3\%$

The maximum smooth Lyapunov exponent: $\widehat{\overline{\lambda}}_1 = -0.36\ (-0.45, -0.20)$

Within the parentheses after the point estimate of $\overline{\lambda}_1$, we report a 95% confidence interval deduced from the Monte Carlo study presented in the next section (see Figure 3), i.e., the quantile of the 1000 simulated copies minus the estimate $\widehat{\overline{\lambda}}_1$ are used. This corresponds to a parametric bootstrap confidence interval (Efron and Tibshirani, 1993).[13]

[11] The choice of these periods are justified in Bask and de Luna (2001), see also Footnote 2.

[12] To compute $\overline{\lambda}_1$ the numerical method described in Eckmann and Ruelle (1985) was used.

[13] Showing that the bootstrap estimation of the distribution of the estimator $\widehat{\overline{\lambda}}$ is consistent is outside the scope of this paper.

No (G)ARCH (generalized autoregressive conditional heteroskedasticity, see, e.g., Shephard, 1996) structure was found by looking at the autocorrelation functions of the squared residuals.

We see that although the model for Period II is less stable than the model for Period I, the difference in stability is not significant when looking at the confidence intervals for the true exponents. On the other hand, we note that the volatility due to exogenous shocks, $Var(\varepsilon_t)$, has decreased by a factor of about 2.3, which is significant using an F-test. This analysis seems to indicate that the decrease in volatility from Period I to Period II is due more to a decrease in exogenous shocks to the economic system than to a change in the stability properties of the model. However, a change in the amplitude of the exogenous shocks will, in general, also have an effect on the stability properties; see the discussion on volatility decomposition of Section 3.2.

In a companion paper (Bask and de Luna, 2001), we study in more depth the properties of exchange rates, and we, for instance, relate the stability properties of European exchange rates against the ECU, the Euro and the U.S. Dollar.

6 Monte Carlo study

In this section, the small sample properties of the presented estimators for $\overline{\lambda}$ are studied. For that purpose, we first consider a polynomial autoregressive model of order one

$$x_t = \theta_1 + \theta_2 x_{t-1} + \theta_3 x_{t-1}^2 + \theta_4 x_{t-1}^3 + v(\theta_1 + \theta_2 x_{t-1} + \theta_3 x_{t-1}^2 + \theta_4 x_{t-1}^3)\varepsilon_t, \tag{11}$$

where $\{\varepsilon_t\}$ is an independent stochastic process with a uniform distribution, $U(-1,1)$, and $v(x) = \min(s, abs(x-b_1), abs(x-b_2))$. Here, s is a scale parameter, and $[b_1, b_2]$ is the basin of attraction of the skeleton of the model. Thus, the distribution of the error term is state dependent to ensure ergodicity (Chan and Tong, 1994).

Three dynamics with different properties are considered, and these are described in Table 1. We set $s = 0.5$ in all cases, which means that we use a stochastic component with a much larger variation than in previous studies (see, e.g., Dechert and Gençay, 1992, McCaffrey et al., 1992, and Whang and Linton, 1999). Our view is that, in many economic applications, the stochastics are large compared with the variability implied by the dynamics. We consider that this is due to the simplicity of the models used in comparison with the complexity of the phenomena under study. The application in the previous section is a typical example.

Table 1: Description of the dynamics

dynamics	$\theta_1, \theta_2, \theta_3, \theta_4$	attractor	$[b_1, b_2]$
(d1)	$0, 0.5, 0.25, 0.1$	limit point	$[-5/4 - 1/4\sqrt{105}, -5/4 + 1/4\sqrt{105}]$
(d2)	$1, 0, -1, 0$	limit cycle	$[-0.5 - 0.5\sqrt{5}, 0.5 + 0.5\sqrt{5}]$
(d3)	$0, 4, -4, 0$	strange	$[0, 1]$

Time series of lengths $n = 100$ and 400 (and also 800 for dynamics (d1)) were simulated[14] and 1000 replications were carried out. Using these time series, the parameters in model (11) were estimated by ordinary least squares (for dynamics (d2) and (d3), θ_4 was set to zero). These estimated parameters were then used to compute an estimate of the smooth Lyapunov exponent in (7) using $m = n$. The results are reported in Table 2.

Table 2: Simulation results: mean, median and standard deviation for the estimated smooth Lyapunov exponents based on 1000 replicates of length n

n	mean $(\widehat{\overline{\lambda}})$	median $(\widehat{\overline{\lambda}})$	s.d. $(\widehat{\overline{\lambda}})$
	d1); $\overline{\lambda} \in [-0.591 \pm 0.002]$		
100	-0.75	-0.72	0.25
400	-0.66	-0.65	0.15
800	-0.63	-0.62	0.11
	d2); $\overline{\lambda} \in [-0.028 \pm 0.001]$		
100	-0.03	-0.03	0.10
400	-0.03	-0.03	0.05
	d3); $\overline{\lambda} \in [0.862 \pm 0.001]$		
100	0.86	0.86	0.06
400	0.86	0.86	0.03

We also provide, in Table 2, a "narrow" confidence interval for the true value of $\overline{\lambda}$ computed as follows:[15] (7) was computed with a simulated time series of length 800 using the known true value of the parameter $\boldsymbol{\theta}_0$. This was repeated 1000 times and a reliable estimate of $\overline{\lambda}$ was obtained by averaging these 1000 replicates. A standard error for $\widehat{\overline{\lambda}}$ was computed from the 1000 replicates, and a 95% confidence interval was calculated based on the normal approximation (see Bailey et al., 1997, for a CLT). This gives a reliable indication of the true value of $\overline{\lambda}$ which we use for comparison. Thus, we see that for dynamics (d2) and (d3), the bias is small and the standard error of the estimated Lyapunov exponent decreases with n as expected. On the other hand, a non-negligible bias is observed in the case of dynamics (d1), although it decreases with n.

Using $m = n$ is justified theoretically for dynamics (d1), where $\frac{d}{dx} f(x; \boldsymbol{\theta}_0) \neq 0$ for all $x \in [b_1, b_2]$. This is not the case for dynamics (d2) and (d3) (see Remark 1), although the simulation results indicate that asymptotic normality is a reasonable approximation even when $m = n$ (see Figure 1). For dynamics (d2), we use $m = 50$, 100 and 200 to study the effect of taking $m < n$ (see Figure 2). The m observations used are chosen randomly out of the n available.[16] As expected, the efficiency decreases with m decreasing, and, although it is not proved theoretically, it seems that $m = n$ is to be preferred. This was also observed in McCaffrey et al. (1992), and Whang and Linton (1999) for non-parametric estimators.

We now turn to polynomial models of order $p > 1$, where the Lyapunov spectrum is estimated with (8). For this purpose, we consider a polynomial autoregression of the form

[14] In all cases, 200 values were first simulated and then discarded.

[15] The value of $\overline{\lambda}$ depends not only on the dynamics but also on the distribution of $v(x_{t-1})\varepsilon_t$. Thus, for instance, although the dynamics of (d2) has a limit cycle, its smooth Lyapunov exponent does not need to be zero.

[16] One may, for instance, also have chosen the observations at regular time intervals.

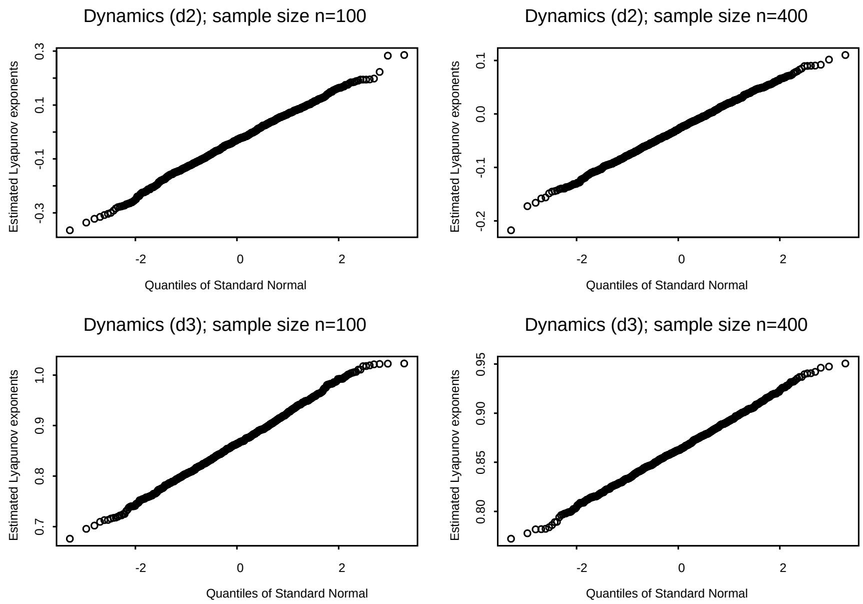

Figure 1: Qq-plots for the 1000 replicates of estimated Lyapunov exponents obtained for dynamics (d2) and (d3) with samples sizes $n = 100$ and 400.

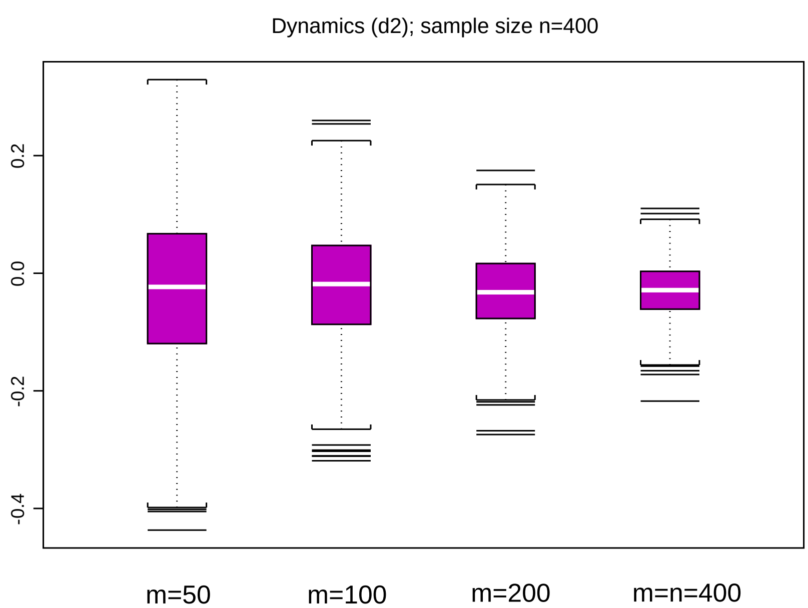

Figure 2: Boxplots for the estimated Lyapunov exponents obtained using different values for m, but based on the same 1000 replicates of size n.

(10) with parameters corresponding to Periods I and II, respectively (see the application in the previous section). The stochastic component is simulated from the normal distribution with mean zero and variances as estimated in Periods I and II. Note that the normal distribution should be truncated to guarantee the stationarity of the process. However, this was found unnecessary in the simulation, i.e., no explosive behavior was observed due to the small variance of the stochastics. The results are summarized in Figure 3 for the maximum smooth Lyapunov exponent, $\widehat{\widehat{\lambda}}_1$. We observe that the estimated Lyapunov exponent has a skewed distribution, and, for Period I, an important bias occurs.

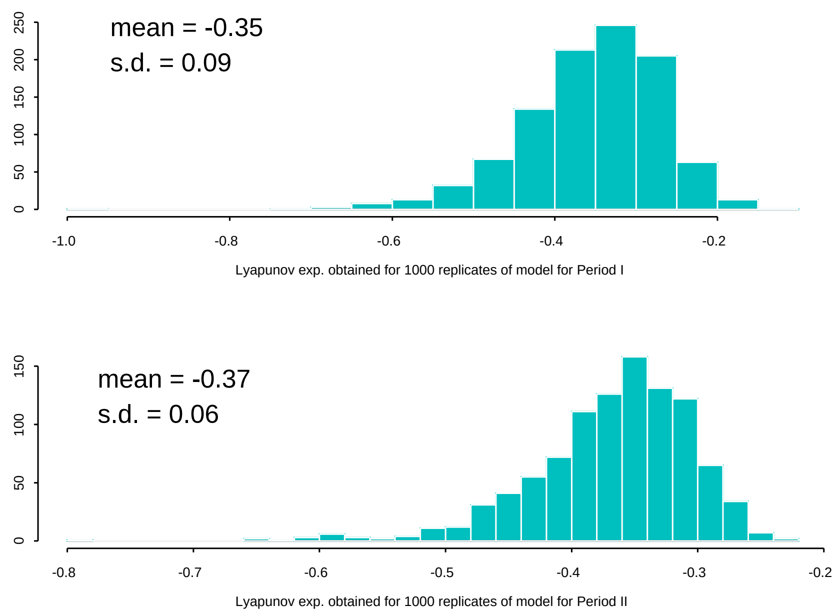

Figure 3: Histograms of the estimated Lyapunov exponents obtained from 1000 replicates simulated with models for Period I and II, respectively.

7 Discussion

Lyapunov exponents have mainly been considered in connection with the investigation of chaotic dynamics. In the econometrics and statistics literature, this has led, for example, Dechert and Gençay (1992) and Nychka et al. (1992a) to consider the estimation of smooth Lyapunov exponents for chaotic systems perturbed by a small amount of noise. Their estimator, which is based on a non-parametric estimation of the dynamics, was further studied in Whang and Linton (1999) who have provided useful asymptotic results. If the purpose is to uncover chaos in a dynamic reconstruction framework (in the spirit of Takens (1981) theorem), then non-parametric regression models are, indeed, most appropriate.

In this paper, smooth Lyapunov exponents have been proposed and studied as a measure of the degree of stability for general dynamic parametric models in order to enhance the interpretability of the latter. As an application, we have investigated the stability of the exchange rates for the Swedish Krona against the ECU and the Euro before and after the implementation of the EMU in January 1999. Our analysis indicated that the decrease in volatility is due more to a decrease in the amplitude of the exogenous shocks to the economic system than to a stabilization of the exchange rate. We have indeed emphasized that the degree of stability of a stochastic dynamic model should be studied when the volatility of a variable modelled is of interest. There is, however, a need for more theoretical work on the relationship between volatility and stability.

In particular, the generalization of Lyapunov type measures to ARCH models (dynamic systems with multiplicative noise) remains an open question. In the presented case study, no ARCH structure was found. This is probably due to the relative short time periods analysed. We have tried to explain changes in volatilities between two periods of time as a result of a change in the dynamics of the system. However, in financial applications ARCH models are often used to explain such volatility changes. Whether a varying volatility is due to changes in the dynamics of the system (structural changes) or to the presence of dynamics in the volatility is a difficult question which cannot be decided by solely looking at the data.

8 Acknowledgments

This paper was presented at the Annual Meeting of the Finnish Statistical Society, 2001, in Vaasa, Finland, at the First Nordic Econometric Meeting in Sönderborg, Denmark, at the NBER/NSF Time Series Workshop 2001 in Raleigh, U.S.A., and at seminars at the Departments of Economics and at the Department of Computing Science, Umeå University, Sweden. Participants at these meetings as well as an anonymous referee are thanked for their valuable comments. Financial support from the Jan Wallanders and Tom Hedelius Foundation is gratefully acknowledged.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.

Bailey, B.A., Ellner, S. and Nychka, D.W. (1997). Chaos with confidence. *Fields Institute Communication*, **11**, 115-133.

Bartlett, M.S. (1950). Periodogram analysis and continuous spectra. *Biometrika*, **37**, 1-16.

Bask, M. and de Luna, X. (2001). EMU and the stability and volatility of foreign exchange: Some empirical evidence. *Umeå Economics Studies*, **565**, Umeå University, Sweden.

Chan, K.S. and Tong, H. (1994). A note on noisy chaos. *Journal of the Royal Statistical Society B*, **56**, 301-311.

Chow, G.C. (1975). *Analysis and control of dynamic economic systems.* New York: John Wiley & Sons.

Dechert, W.D. and Gençay, R. (1992). Lyapunov exponents as a nonparametric diagnostic for stability analysis. *Journal of Applied Econometrics*, **7**, S41-S60.

de Luna, X. (1998). Projected polynomial autoregression for prediction of stationary time series. *Journal of Applied Statistics*, **25**, 763-775.

Eckmann, J.-P. and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, **57**, 617-656.

Efron, B. and Tibshirani, R.J. (1993). *An introduction to the bootstrap.* New York: Chapman and Hall.

Granger, C.W.J. and Teräsvirta, T. (1993). *Modelling nonlinear economic relationships.* Oxford: Oxford University Press.

Herzel, H., Ebeling, W. and Schulmeister, T. (1987). Nonuniform chaotic dynamics and effects of noise in biochemical systems. *Zeitschrift für Naturforschung*, **42a**, 136-142.

McCaffrey, D., Ellner, S., Gallant, A.R. and Nychka, D. (1992). Estimating the Lyapunov exponent of a chaotic system with nonparametric regression. *Journal of the American Statistical Association*, **87**, 682-695.

Nychka, D., Ellner, S., Gallant, A.R. and McCaffrey, D. (1992a). Finding chaos in nonlinear systems. *Journal of the Royal Statistical Society B*, **54**, 399-426.

Nychka, D., Ellner, S., Gallant, A.R. and McCaffrey, D. (1992b). Discussion on the meeting on chaos. *Journal of the Royal Statistical Society B*, **54**, 469-471.

Plosser, C.I. (1989). Understanding real business cycles. *Journal of Economic Perspectives*, **3**, 51-57.

Potter, S.M. (2000). Nonlinear impulse response functions. *Journal of Economic Dynamics and Control*, **24**, 1425-1446.

Ruelle, D. (1989). *Chaotic evolution and strange attractors.* Cambridge: Cambridge University Press.

Shephard, N.G. (1996). Statistical aspects of ARCH and stochastic volatility. In *Time Series Models in Econometrics, Finance and other fields* by Cox, D.R., Hinkley, D.V. and Barndorff-Nielsen, O.E., eds., London: Chapman and Hall, 1-67.

Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence* by Rand, D. and Young, L., eds., Berlin: Springer-Verlag, 366-381.

Tong, H. (1995). A personal overview of non-linear time series analysis from a chaos perspective (with discussion). *Scandinavian Journal of Statistics*, **22**, 399-445.

Whang, Y.-J. and Linton, O. (1999). The asymptotic distribution of nonparametric estimates of the Lyapunov exponent for stochastic time series. *Journal of Econometrics*, **91**, 1-42.

White, H. and Domowitz, I. (1984). Nonlinear regression with dependent observations. *Econometrica*, **52**, 143-161.

Wolff, R.C.L. (1992). Local Lyapunov exponent: Looking closely at chaos. *Journal of the Royal Statistical Society B*, **54**, 353-372.

Yao, Q. and Tong, H. (1994). Quantifying the influence of initial values on non-linear prediction. *Journal of the Royal Statistical Society B*, **56**, 701-725.