Post-Sample Prediction Tests for the Efficient Method of Moments

Pieter J. van der Sluis*
Department of Actuarial Science and Econometrics
University of Amsterdam
and
Tinbergen Institute

April 21, 1997

Abstract

In this paper a post-sample prediction test is derived for estimators based on the Efficient Method of Moments. The main advantage of this particular test over other stability tests is that no time-consuming estimation of the structural parameters for the post-sample is needed. The asymptotic properties of the test and local power properties against certain alternatives are deduced. Using the Efficient Method of Moments methodology, an application is made to stochastic volatility models for the British pound versus Canadian dollar exchange rates. The breakpoint for the stability test is a priori set at September 16th 1992, when Britain was forced to leave the European Monetary Union Exchange Rate System.

Keywords: Efficient method of moments; Exchange rates; Specification testing; Stochastic volatility models.

JEL-classification: C15, C52

1 Introduction

In case the likelihood of a model is intractable one may resort to simulation techniques such as *indirect inference (*Gourieroux *et al*. (1993)) and the Efficient Method of Moments (EMM) (Gallant and Tauchen, (1996)). Both

*Address for correspondence: Pieter J. van der Sluis, University of Amsterdam, Department of Actuarial Science and Econometrics, Roetersstraat 11, NL-1018 WB Amsterdam, The Netherlands, phone +31 20 525 4257, fax +31 20 525 4349, email Sluis@butler.fee.uva.nl.
A previous version of this paper entitled *Post-sample prediction tests for indirect inference estimators* was presented at the sixth meeting of EC^2, 14-16 December 1995 at the University of Aarhus, Aarhus, Denmark and at the North American Summer Meeting of the Econometric Society, University of Iowa, Iowa, USA, 21-24 June 1996. The author thanks Peter Boswijk, Ruud Koning, Jesper Lund, Harold Zhang and participants at the above-mentioned conferences for comments. The author thanks Jurgen Doornik for feedback on issues on the Ox programming language. Financial support from the Tinbergen Institute is gratefully acknowledged.

methods may be seen as *minimum chi square* estimators. Therefore the *post-sample prediction* (PSP) statistic for GMM estimators, as given in Hoffman and Pagan (1989) and Ghysels and Hall (1990a), may be transcribed to these type of simulation-based estimators. In this paper we will do so for the EMM estimation technique. We believe that EMM is superior to the original indirect inference technique of Gourieroux *et al*. (1993) from a viewpoint of efficiency and stability; see Van der Sluis (1996) for details. Although the results in this paper can easily be applied to the *indirect inference* case, it is doubtful whether this will ever be necessary for substantial models. Because of the lack of efficiency of indirect inference.

The PSP test is based on an evaluation of the moment conditions at a sample-based parameter estimate and data from the post-sample. Like the PSP test for GMM, the PSP test for EMM may detect parameter instability. The test may be used in cases where there is prior knowledge about switches in regime or as a specification test for sample and post-sample chosen of the same size. Whether the latter procedure provides reliable conclusions is questionable in the light of recent results of Hall and Sen (1996) for the GMM case. The underlying principle of this test is different from that of the LM, Wald or LR based statistics for structural stability for GMM estimators as developed by Andrews and Fair (1988). In Ghysels and Hall (1990a) it is argued that the PSP test has several advantages over these Wald, LR and LM based statistics. Three advantages are mentioned, for which the first one is even more important for simulation based estimators than for GMM estimators: (i) we do not need an estimate of the structural parameter over the post-sample; (ii) all orthogonality restrictions are used over the post-sample; (iii) no subset of orthogonality restrictions is set equal to zero in the post-sample. The latter two advantages are reflected in the number of degrees of freedom and its local power properties against certain alternatives.

This paper is organised in the following manner. Section 2 describes the EMM method as proposed by Gallant and Tauchen (1996). This method often provides an algorithm that is computationally more stable and attractive than the indirect inference procedure. Moreover, this method will provide full maximum-likelihood efficiency of the estimators, see Gallant and Tauchen (1996), Gallant and Long (1996) and Tauchen (1996). This section is merely intended to provide some notation. The reader is advised to consult the original paper in case of nonfamiliarity with EMM. In section 3 the PSP statistic for GMM is discussed. In section 4 the PSP statistics for EMM are proposed. Its asymptotic properties and local power against certain alternatives are deduced in section 5. In this paper applications will be made in the context of *stochastic volatility* models. Stochastic volatility models are models for which the likelihood has no tractable expression. These models are mainly used to describe the heteroskedasticity process of financial time-series, replacing ARCH-type models. The implementation of the test in the context of stochastic volatility models will be considered in section 6. In this section the choice of the auxiliary model, which is needed in EMM, is motivated. Also the problem of determining the number of moment conditions for sample and post-sample is assessed. In section 7 exchange rates are investigated using these stochastic volatility models. Parameter instability is associated with periods of different volatility. Section 8 concludes.

2 Efficient Method of Moments

Indirect inference techniques are based on some connection between the auxiliary model and the structural model. In the original indirect inference procedure of Gourieroux *et al*. (1993) this link is established through the parameters of the auxiliary model (*parameter calibration*). No strict guidelines are given for the choice of the auxiliary model. Gallant and Tauchen (1996) propose the *efficient method of moments* (EMM). Here the connection between the auxiliary model and the dynamic model is achieved by means of the expected scores of the auxiliary model under the structural model (*score calibration*). Since scores are better comparable over different models than parameters, this will typically yield more stable optimization problems. More importantly, strict guidelines are given for the choice of the auxiliary model such that maximum likelihood efficiency is attained. Another advantage is that the score has to be evaluated only once per optimization round, whereas for indirect inference per optimization round a dozen of auxiliary models has to be estimated.

In short, EMM goes as follows. Consider the structural model (1) below. Several regularity conditions must be met by the dynamic model and the auxiliary model. For details see Gallant and Tauchen (1996). Here it suffices to note that the original results in Gallant and Tauchen (1996) were presented for *Markovian* models, however recent results of Gallant and Long (1996) justify the use of *non-Markovian* auxiliary models. However, the EMM framework is up till now limited to *stationary* and *ergodic* structural models. The structural model is defined as

$$\begin{aligned} y_t &= r(y_{t-1}, x_t, u_t, \theta) \\ u_t &= \phi(u_{t-1}, \epsilon_t, \theta) \end{aligned} \tag{1}$$

where θ is a parameter, $\theta \in \Theta \subseteq \mathbb{R}^p$. The sequence of densities for the structural model (1) will be denoted

$$\{p_1(x_1 \mid \theta), \{p_t(y_t \mid x_t, \theta)\}_{t:=1}^{\infty}\}$$

The sequence of densities for the auxiliary process will be denoted as

$$\{f_1(w_1 \mid \beta), \{f_t(y_t \mid w_t, \beta)\}_{t:=1}^{\infty}\}$$

where x_t and w_t are observable endogenous variables. Let us define

$$m(\theta, \beta) := \int \int \frac{\partial}{\partial \beta} \ln f(y \mid w, \beta) p(y \mid x, \theta) dy p(x \mid \theta) dx$$

the expected score of the auxiliary model under the dynamic model. The simulation approach solely consists of calculating this function as

$$m_N(\theta, \beta) := \frac{1}{N} \sum_{\tau:=1}^{N} \frac{\partial}{\partial \beta} \ln f(y_\tau(\theta) \mid w_\tau(\theta), \beta)$$

for N very large. The EMM estimator $\widehat{\theta}_n(\mathcal{I}_n)$ is defined as

$$\widehat{\theta}_n(\mathcal{I}_n) := \arg \min_{\theta \in \Theta} m_N^{'}(\theta, \widehat{\beta}_n)(\mathcal{I}_n)^{-1} m_N(\theta, \widehat{\beta}_n)$$

where $\mathcal{I}_n$ is a weighting matrix and $\widehat{\beta}_n$ is a consistent estimator of the parameters in the auxiliary model.

In the notation of this section we have

$$\mathcal{I}_n = V_0[\frac{1}{\sqrt{n}} \sum_{t:=1}^{n} (\frac{\partial}{\partial \beta} \ln f_t(y_t \mid w_t, \widehat{\beta}_n)]$$

and obviously

$$\lim_{n \to \infty} \mathcal{I}_n = \mathcal{I}_0$$

One can also prove for the estimates for the structural model

$$\sqrt{n} m_N(\theta_0, \widehat{\beta}_n) \xrightarrow{d} N(0, \mathcal{I}_0)$$

and consistency of θ :

$$\sqrt{n}(\widehat{\theta}_n(\mathcal{I}_0) - \theta_0) \xrightarrow{d} N(0, [\mathcal{M}_0^{'}(\mathcal{I}_0)^{-1}\mathcal{M}_0]^{-1})$$

where $\mathcal{M}_0 := \frac{\partial}{\partial \theta'} m(\theta_0, \beta^*)$. For notational convenience $\widehat{\theta}_n$ will denote the estimator based on the optimal $\mathcal{I}_n$.

In order to justify the full efficiency of EMM, one needs to employ for example the SemiNonParametric (SNP) density of Gallant and Nychka (1987)[1]. Gallant and Long (1996) show that if we use a SNP approach to select the auxiliary model, the full efficiency can be achieved when the SNP model is rich enough such that the structural model lies in the SNP hierarchy. To improve the finite sample properties of EMM one should add a parametric part to the SNP density. The polynomial part of the SNP density will now be relieved from some of its tasks. In the literature this parametric part is referred to as the leading term of the SNP expansion. The process $y_t(\theta_0)$ is the process under investigation, $\mu_t(\theta_0) := \mathsf{E}_{t-1}[y_t(\theta_0)]$, is the conditional mean, $\epsilon_t(\theta_0) := y_t(\theta_0) - \mu_t(\theta_0)$, $\sigma_t^2(\theta_0) := \mathsf{Var}_{t-1}[\epsilon_t(\theta_0)] = \mathsf{E}_{t-1}[\epsilon_t^2(\theta_0)]$ is the conditional variance and $z_t(\theta_0) := \epsilon_t(\theta_0)(\sigma_t^2(\theta_0))^{-1/2}$ is the standardized process. The SNP density now takes the following form

$$f(z_t(\theta)) = \frac{1}{\sqrt{\sigma_t^2}} \frac{[P_K(z_t, x_t)]^2 \phi(z_t)}{\int [P_K(u, x_t)]^2 \phi(u) du}$$

where ϕ denotes the standard normal density, $x := (y_{t-1}, ..., y_{t-M})$ and the polynomials

$$P_K(z, x_t) := \sum_{i:=0}^{K_z} a_i(x_t) z^i := \sum_{i:=0}^{K_z} [\sum_{j:=0}^{K_x} a_{ij} x_t^j] z^i$$

For identification we set $a_{00} = 1$. A specific form for the polynomials is taken, viz. orthogonal Hermite polynomials (see Gallant *et al.* (1991), Fenton and Gallant (1996a) and Andersen and Lund (1996)). Relevant formulas can be found in Abramowitz and Stegun (1972) and Fenton and Gallant (1996a). For an assessment of the properties of SNP density estimators see Fenton and Gallant (1996a+b). In Gallant *et al.* (1991) and Tauchen (1996) a very general SNP hierarchy of models is considered. Since in our set-up x_t is a vector of lagged values of y_t the polynomials in terms of x_t become so-called *monomials* for $M > 1$. In this case the analysis becomes much more complicated and hardly relevant with a proper parametric leading term. Therefore in this paper only the case $M = 0$ and $M = 1$ are considered.

In this paper structural stability is of interest, therefore the following nomenclature is introduced: n_1 will denote the number of observation in the sample and n_2 will denote the number of observations in the post-sample. The auxiliary estimator that employs sample data only, will be denoted $\widehat{\beta}_{n_1}$. The estimator that employs post-sample data only will be denoted $\widehat{\beta}_{n_2}$. Likewise the EMM estimators $\widehat{\theta}_{n_1}$ and $\widehat{\theta}_{n_2}$ are denoted. Furthermore estimators that are based on both sample and post-sample data are denoted $\widehat{\beta}_{n_1+n_2}$ and $\widehat{\theta}_{n_1+n_2}$. In this paper we will take $p := \dim(\theta)$, $q_1 := \dim(\beta_1)$ and $q_2 := \dim(\beta_2)$. The number of moment conditions q_1 and q_2 will be determined using several criteria. For important results regarding EMM to hold it is necessary that the number of moment conditions in the sample q_1 increases with n_1 and the number of moment conditions in the post-sample q_2 increases with n_2. Note in this respect the conceptual difference with GMM. It will automatically happen that q_1 and q_2 will increase with n_1 and n_2, respectively, using any of the criteria such as the Akaike Information Criterion (AIC, Akaike (1973)), the Schwarz Criterion (BIC, Schwarz (1978)) or the Hannan-Quinn Criterion (HQC, Hannan and Quinn (1979) and Quinn (1980))[2]. In section 4 we will see that the power properties depend on the number and type of moment conditions. Some guidelines will be given for optimal results in practice.

[1] Building on earlier work of Phillips (1983)

[2] AIC= $T \ln L + |\beta|$

BIC= $T \ln L + 0.5|\beta| \ln T$

3 PSP-statistic for GMM

Hoffman and Pagan (1989) and Ghysels and Hall (1990a) constitute for GMM an analogy to OLS where the average prediction error over a post-sample prediction period of $n_1+1, ..., n_1+n_2$, i.e.

$$n_2^{-1}\sum_{i:=n_1+1}^{n_1+n_2}\widehat{e}_i,$$

is sometimes considered[3]; here $\widehat{e}_i$ denotes the standard OLS residual evaluated at $\widehat{\theta}_{n_1}$ For GMM this becomes

$$\widehat{\tau} := n_2^{-1}\sum_{i:=n_1+1}^{n_1+n_2}\widehat{g}_i$$

where $\widehat{g}_i$ is a vector of first order conditions g evaluated at $\widehat{\theta}_{n_1}$. This test may thus indicate parameter instability over the post-sample area. Let $T_1 := \{1,\ldots,n_1\}$ and $T_2 := \{n_1+1,\ldots,n_1+n_2\}$. In Ghysels and Hall (1990a) the asymptotic distribution of $\sqrt{n_2}\widehat{\tau}$ is given. It turns out that $\sqrt{n_2}\widehat{\tau}$ is asymptotically normal:

$$\sqrt{n_2}\widehat{\tau} \xrightarrow{d} N(0, S_2 + k\widetilde{D}_2(D_1^{'}S_1^{-1}D_1)^{-1}\widetilde{D}_2^{'})$$

where $S_i := \lim_{n_i\to\infty} \mathsf{E}[n_i g_{n_i} g_{n_i}^{'}]$, $D_i = \text{plim}_{n_i\to\infty} \frac{\partial g_{n_i}}{\partial\theta'}$ and $\widetilde{D}_2 := \text{plim}_{n_2\to\infty} \frac{\partial g_{n_i}(\widehat{\theta}_{n_1})}{\partial\theta'}$, $i = 1, 2$. Under stationarity we have $S_1 = S_2$.

Following Newey (1985), Ghysels and Hall (1990b) and Ahn (1995) the following individual null hypotheses are identified in the context of structural stability and GMM estimation

$$\begin{aligned} \mathsf{H}_0^1 &: \mathsf{E}f_{n_1}(y_t,\theta_0) = 0,\ \forall\, t\in T_1 \\ \mathsf{H}_0^2 &: \mathsf{E}f_{n_2}(y_t,\overline{\theta}_0) = 0,\ \forall\, t\in T_2 \\ \mathsf{H}_0^3 &: \theta_0 = \overline{\theta}_0 \end{aligned}$$

where the function f represents the moment restrictions. The subscripts n_1 and n_2 refer to the sample and post-sample respectively. In case of using this test as a general stability test the role of sample and post-sample may be reversed.

The *stability* hypothesis H_0^s is defined as

$$\mathsf{H}_0^s : \mathsf{H}_0^1, \mathsf{H}_0^2 \text{ and } \mathsf{H}_0^3 \text{ hold}$$

The asymptotic power of the test depends on which of the individual null-hypotheses is violated. Define the following alternative hypotheses

$$\begin{aligned} \mathsf{H}_A^A &: \text{At least one of } \mathsf{H}_0^1, \mathsf{H}_0^2 \text{ and } \mathsf{H}_0^3 \text{ does not hold} \\ \mathsf{H}_A^B &: \mathsf{H}_0^1 \text{ holds, at least one of } \mathsf{H}_0^2 \text{ and } \mathsf{H}_0^3 \text{ does not hold} \\ \mathsf{H}_A^C &: \mathsf{H}_0^1 \text{ and } \mathsf{H}_0^2 \text{ hold, } \mathsf{H}_0^3 \text{ does not hold} \end{aligned}$$

HQC$= T\ln L + |\beta|\ln\ln T$

In this context, L denotes the loglikelihood of the auxiliary model in the optimum, T denotes the number of observations and $|\beta|$ the number of freely estimated auxiliary parameters. For some properties see Lütkepohl (1990).

[3] Another analogy may be constituted by $n_2^{-1}\sum_{i:=n_1+1}^{n_1+n_2} x_i\widehat{e}_i$

This leads to the following local alternatives

$$\begin{aligned}
\mathsf{H}_n^1 &: \sqrt{n_1}\mathsf{E} f_{n_1}(y_t,\theta_0) = \delta_1 + o(1) \\
\mathsf{H}_n^2 &: \sqrt{n_2}\mathsf{E} f_{n_2}(y_t,\theta_0) = \delta_2 + o(1) \\
\mathsf{H}_n^3 &: \mathsf{E} f_{n_2}(y_t,\overline{\theta}_{0,n}) = 0,\ \sqrt{n_2}(\overline{\theta}_{0,n} - \theta_0) = \delta
\end{aligned}$$

Let $\mathsf{H}_l^A := \{(\mathsf{H}_n^1,\mathsf{H}_n^2)\}_{n:=1}^{\infty}$, $\mathsf{H}_l^B := \{(\mathsf{H}_0^1,\mathsf{H}_n^2)\}_{n:=1}^{\infty}$ and $\mathsf{H}_l^C := \{(\mathsf{H}_0^1,\mathsf{H}_n^3)\}_{n:=1}^{\infty}$. The usual Hansen $J-$test may be used to test H_0^s. However, Ghysels and Hall (1990b) show that this test has low power against local alternatives that are of the form H_l^C. In Ghysels and Hall (1990a) a modification of the Hansen test is proposed.

It should be noted that recently Hall and Sen (1996) and Sowell (1996) proposed a different decomposition of the stability hypothesis for GMM. However, considering the different nature of the moment conditions between GMM and Indirect Inference/EMM, the present decomposition will be used in the next sections to construct a stability hypothesis in the context of simulation techniques.

Ahn (1995) shows, building on earlier work of Newey (1985) that the PSP test for GMM is an optimal GMM test that has maximum power toward H_l^B. Several interesting asymptotic equivalencies are described for the tests for structural stability proposed by Andrews and Fair (1988). The upshot is that any of their statistics has maximum local power against alternatives of the form H_l^C. In the context of EMM these statistics are investigated in an accompanying paper Van der Sluis (1996).

4 PSP-statistic for EMM

Define $\beta_i = \text{plim}_{n_i\to\infty}\widehat{\beta}_{n_i}$ then for the EMM case the following individual null hypotheses are identified

$$\begin{aligned}
\mathsf{H}_0^1 &: m(\theta_0,\beta_1) = 0,\ \forall\, t \in T_1 \\
\mathsf{H}_0^2 &: m(\overline{\theta}_0,\beta_2) = 0,\ \forall\, t \in T_2 \\
\mathsf{H}_0^3 &: \theta_0 = \overline{\theta}_0
\end{aligned}$$

where the function m is the expected score of the auxiliary model under the dynamic model as defined in the previous section. The *stability* hypothesis and alternative hypotheses are as in the previous section.

Let $\widehat{\theta}_{n_1+n_2}$ denote a root-n consistent estimator of the combined sample and post-sample. The following local alternatives are defined for EMM

$$\begin{aligned}
\mathsf{H}_n^1 &: \sqrt{n_1} m(\theta_0,\beta_1) = \delta_1 + o(1) \\
\mathsf{H}_n^2 &: \sqrt{n_2} m(\theta_0,\beta_2) = \delta_2 + o(1) \\
\mathsf{H}_n^3 &: m(\overline{\theta}_{0,n},\beta_2) = 0,\ \sqrt{n_2}(\overline{\theta}_{0,n} - \theta_0) = \delta
\end{aligned}$$

Let $\mathsf{H}_l^A := \{(\mathsf{H}_n^1,\mathsf{H}_n^2)\}_{n:=1}^{\infty}$, $\mathsf{H}_l^B := \{(\mathsf{H}_0^1,\mathsf{H}_n^2)\}_{n:=1}^{\infty}$ and $\mathsf{H}_l^C := \{(\mathsf{H}_0^1,\mathsf{H}_n^3)\}_{n:=1}^{\infty}$. In an accompanying paper, Van der Sluis (1996), different types of stability tests are investigated. These tests have optimal asymptotic power against H_l^A and H_l^C. In this paper we only consider local alternatives of the form H_l^B. Since the PSP test in the GMM case is optimal against H_l^B, it is tempting to propose a PSP-statistic for EMM based on $m_N(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})$. In the next section the asymptotic properties of the statistics based on this quantity will be deduced.

5 Asymptotic Properties and Local Power

In this section some theorems on the asymptotic distributions and local power properties of the PSP-test will be provided. The results in this section are basically transcriptions from results known in the GMM literature to the realm of EMM. A proof of the following theorem can be found in the appendix.

Theorem 1 *Let*

$$\mathsf{PSP}_{\mathsf{EMM}} := n_2 m_N^{'}(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})\{\widehat{\mathcal{I}}_{n_2} + k\mathcal{M}_2(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})[\mathcal{M}_1^{'}(\widehat{\theta}_{n_1},\widehat{\beta}_{n_1})\widehat{\mathcal{I}}_{n_1}^{-1}\mathcal{M}_1(\widehat{\theta}_{n_1},\widehat{\beta}_{n_1})]^{-1} \cdot \mathcal{M}_2^{'}(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})\}^{-1} m_N(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})$$

under H_l^B $\mathsf{PSP}_{\mathsf{EMM}}$ *has a noncentral* $\chi^2_{q_2}$ *distribution with noncentrality parameter* $\lambda_{\mathsf{PSP}_{\mathsf{EMM}}}$*, given by*

$$\lambda_{\mathsf{PSP}_{\mathsf{EMM}}} := \delta_2^{'}\mathcal{I}_2^{-1}[\mathcal{I}_2 - c_2\mathcal{M}_2(\theta_0,\beta_2)[\sum_{i:=1}^{2}\mathcal{M}_i^{'}(\theta_0,\beta_i)\mathcal{I}_i^{-1}\mathcal{M}_i(\theta_0,\beta_i)]^{-1}\mathcal{M}_2^{'}(\theta_0,\beta_2)]\mathcal{I}_2^{-1}\delta_2$$

where $c_2 := \lim_{n_1,n_2\to\infty}\frac{n_2}{n_1+n_2}$.

Of course the role of the sample and post-sample may be reversed, this leads to a *backward post-sample prediction* or *ante-sample prediction* test, $\overleftarrow{\mathsf{PSP}}_{\mathsf{EMM}}$ test.

Theorem 2 *Let*

$$\overleftarrow{\mathsf{PSP}}_{\mathsf{EMM}} := n_1 m_N^{'}(\widehat{\theta}_{n_2},\widehat{\beta}_{n_1})\{\widehat{\mathcal{I}}_{n_1} + k^{-1}\mathcal{M}_1(\widehat{\theta}_{n_2},\widehat{\beta}_{n_1})[\mathcal{M}_2^{'}(\widehat{\theta}_{n_2},\widehat{\beta}_{n_2})\widehat{\mathcal{I}}_{n_2}^{-1}\mathcal{M}_2(\widehat{\theta}_{n_2},\widehat{\beta}_{n_2})]^{-1} \cdot \mathcal{M}_1^{'}(\widehat{\theta}_{n_2},\widehat{\beta}_{n_1})\}^{-1} m_N(\widehat{\theta}_{n_2},\widehat{\beta}_{n_1})$$

under $\widetilde{\mathsf{H}}_l^B := \{(\mathsf{H}_n^1,\mathsf{H}_0^2)\}_{n:=1}^{\infty}$ $\overleftarrow{\mathsf{PSP}}_{\mathsf{EMM}}$ *has a noncentral* $\chi^2_{q_1}$ *distribution with noncentrality parameter* $\lambda_{\overleftarrow{\mathsf{PSP}}_{\mathsf{EMM}}}$*, given by*

$$\lambda_{\overleftarrow{\mathsf{PSP}}_{\mathsf{EMM}}} := \delta_1^{'}\mathcal{I}_1^{-1}[\mathcal{I}_1 - c_1\mathcal{M}_1(\theta_0,\beta_1)[\sum_{i:=1}^{2}\mathcal{M}_i^{'}(\theta_0,\beta_i)\mathcal{I}_i^{-1}\mathcal{M}_i(\theta_0,\beta_i)]^{-1}\mathcal{M}_1^{'}(\theta_0,\beta_1)]\mathcal{I}_1^{-1}\delta_1$$

where $c_1 := \lim_{n_1,n_2\to\infty}\frac{n_1}{n_1+n_2}$.

A different role is played by the moment conditions in GMM than EMM and indirect inference: for estimation purposes it may be desirable to vary the number of moment conditions with the size of the sample and post-sample. This will be done automatically by a criterion such as the BIC. Particularly, a high dimensional auxiliary which is needed for EMM to work, may not be of use for small post-samples. For the post-sample the high-dimensional auxiliary model may not converge at all. Note however that changing the auxiliary model for the post-sample induces different null and alternative hypotheses. In Van der Sluis (1996b) it is shown that in the case of LM/LR/Wald/ tests for structural stability it is optimal[4] to change the number of moment conditions. To the author's knowledge, in case of PSP tests no such property can be established. This leads to a choice of the same moment conditions in the post-sample as in the sample. Other optimality properties of the PSP test-statistics follow from straightforward application of results known in the GMM literature, see Ahn (1995) and Newey (1985).

It should be noted that Tauchen (1996) shows that certain misspecifications of the structural model can lead to the failure to reject the overidentifying restrictions. The same possibility can also lead to the failure to reject the stability hypothesis.

[4] *Optimal* in the sense of yielding a uniformly most powerful test

6 Implementation of the test

Recently it is reported that *stochastic volatility models* may provide better estimates for the volatility parameters than the ARCH-GARCH class of models. A nice reference on stochastic volatility models is Ghysels *et al.* (1996). Examples of the estimation of stochastic volatility models by EMM can be found in Gallant *et al.* (1994) and Andersen and Lund (1995). The stochastic volatility model in its ASARMAV(p,q)[5] form reads

$$
\begin{aligned}
&y_t = \sigma_t \epsilon_t \\
&\ln \sigma_t^2 = \omega + \textstyle\sum_{i=1}^{p} \rho_i \ln \sigma_{t-i}^2 + \eta_t + \sigma_\eta (1 + \sum_{i=1}^{q} \zeta_j L^j)\eta_t \\
&\epsilon_t \sim \mathsf{NID}(0,1), \eta_t \sim \mathsf{NID}(0,1), \mathsf{Corr}(\epsilon_t, \eta_{t+1}) = \lambda, \ -1 \le \lambda \le 1 \\
&t = 1, \ldots, T
\end{aligned}
$$

Just as for ordinary ARMA models for covariance-stationarity the roots of $1 - \rho_1 z - \ldots - \rho_p z^p$ must lie outside the unit circle. Since the latent variables σ_t must be integrated out, the log-likelihood of this model cannot be written down in a tractable form. Using EMM we may efficiently estimate the parameters $\theta := \{\omega, \rho_1, \ldots, \rho_p, \zeta_1, \ldots, \zeta_q, \sigma_\eta, \lambda\}$ of this model. This may be done with EmmPack (Van der Sluis (1997)).

As mentioned in section 2 the auxiliary model is taken from the class of SemiNonParametric (SNP) models introduced by Phillips (1983) and later developed by Gallant and Nychka (1987). To relieve the SNP density from its task, a parametric model is added to the SNP density. There is reported evidence that this will improve the finite sample performance of EMM. Our candidate parametric model is taken from the ARCH class of models. We will successively consider two branches of the ARCH class of models: the ARCH-GARCH (Engle (1982), Bollerslev (1986)) models and the EGARCH models (Nelson (1991)).

The ARCH-GARCH class of models reads

$$
\begin{aligned}
&y_t = \sigma_t \epsilon_t \\
&\sigma_t^2 = \alpha_0 + \textstyle\sum_{i:=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j:=1}^{p} \gamma_j \sigma_{t-j}^2 \\
&\epsilon_t \sim \mathsf{NID}(0,1) \\
&t = 1, \ldots, T
\end{aligned}
$$

This model is usually called GARCH(p,q) model. If we set $p = 0$, we obtain a pure ARCH(q) model as introduced in Engle (1982). These models may be used as auxiliary models in the applications of the next sections. For the ARCH(q) model the restrictions on the parameters are $\alpha_0 > 0$ and $\alpha_i \ge 0$ for $i > 0$ for the model to be well-defined and $\sum_{i:=1}^{q} \alpha_i \le 1$ for covariance stationarity. Experimenting with these type of models show that in order to capture all heteroskedasticity in the data the number of parameters in the pure ARCH models gets rather high. This may cause the same problems as occur with estimating stochastic volatility models with GMM using a high number of moment conditions as reported in Andersen and Sørensen (1996). One suggestion may be to use GARCH models. However a practical problem arises. For the GARCH(p,q) model to be well defined we need that all the coefficients in the infinite power series expansion for $\alpha(x)/(1-\gamma(x))$ are non-negative. Here $\alpha(L)$ and $\gamma(L)$ are the usual polynomials with L the lag-operator and provided $\alpha(L)$ and $\gamma(L)$ have no common roots and that the roots of the polynomial $\gamma(x) = 1$ lie outside the unit circle. The model is covariance stationary if and only if all the roots of $\alpha(x) + \gamma(x)$ lie outside the unit circle. Various other stationarity concepts and restrictions are discussed in Kleibergen and Van Dijk (1993) or Nelson and Cao (1992). References can be found in Bollerslev *et al.* (1994). The problem is that in practice often estimates are obtained that fall

[5] Asymmetric Stochastic AutoRegressive Moving Average Volatility

into the strict stationarity region, but outside the covariance stationarity region. This is not a problem per se, but for EMM it is, since EMM hinges on concepts as ergodicity and stationarity, it must be guaranteed that the simulated series generates a stable score-generator. For parameter values outside the most stringent parameter space this cannot be guaranteed. For a more extensive discussion of this issue, reference is made to Andersen and Lund (1996a) and Van der Sluis (1996). Moreover, for GARCH models parameter estimates will often turn out to be on this most stringent IGARCH boundary. On this IGARCH boundary the error variances become infinite. Because of this property a comparison of a long simulated IGARCH series with real financial data shows that IGARCH models are rather far from realistic.

Consequently, the EGARCH models of Nelson (1991) are considered[6]. The EGARCH(p,q) model reads

$$
\begin{aligned}
&y_t = \sigma_t z_t \\
&\sigma_t^2 = \exp\{\alpha_0 + \textstyle\sum_{j:=1}^{p} \gamma_j \ln \sigma_{t-j}^2 + (1 + \sum_{i:=1}^{q} \alpha_i L^i)(\kappa_1 z_{t-1} + \kappa_2[|z_{t-1}| - E|z_{t-1}|])\} \\
&z_t \sim \mathsf{NID}(0,1) \\
&t = 1, \ldots, T
\end{aligned}
$$

The stationarity restrictions for the pure EGARCH can be derived in a similar way to those in ARMA models à la Box-Jenkins. Therefore only the γ parameters play a role in the stationarity restrictions: the model is stationary (*strict and weak*) in case the roots of $\gamma(x)$ lie outside the unit circle. Thus the stationarity problems that arise in GARCH models vanish for EGARCH models

In the next section an application to exchange rates will be made. The class of auxiliary models that will be considered is the EGARCH(p,q)-H(K_x, K_z) class. The order of the auxiliary model will be determined using several criteria, as BIC, AIC and HQC. The vector of auxiliary parameters β has dimension $p + q + 3 + (K_x + 1)(K_z + 1)$. The moment conditions all represent a salient feature of the model. For the EGARCH model we have in particular that the moment conditions associated with the γ parameters give us a idea of the persistence in the data. In particular, $\sum_{i=1}^{p} \gamma_i$ will serve as a measure of persistence. The α_0 parameter is a measure of the unconditional variance. The term $1 + \sum_{i=1}^{q} \alpha_i$ reflects the short-run sensitivity to innovations. The term κ_1 reflects the asymmetric volatility effect that is often present in financial time series. The κ_2 parameter measures the changes in z_t in reaction with the conditional heteroskedasticity. In case $K_x = 0$, letting $K_z > 0$ induces a time-homogeneous non-Gaussian error structure. The case $K_x > 0$ induces heterogeneous innovation densities beyond the EGARCH model. Since the EGARCH model captures all this heterogeneity, $K_x > 0$ will not be necessary. To the author's knowledge there are no guidelines provided yet for assessment of the individual coefficients of the Hermite polynomial.

Starting values may be obtained by using the *quasi maximum likelihood technique* of Harvey *et al.* (1993). The following was done[7]

(i) Determine the order of the ARCH process. For this particular ARCH process estimate $\widehat{\beta}$. Set the corresponding weighting matrix $\widehat{\mathcal{I}}_n$ equal to $\frac{1}{n}\sum_{t:=1}^{n}[\frac{\partial}{\partial\beta}\ln f_t(y_t|w_t,\widehat{\beta}_n)][\frac{\partial}{\partial\beta}\ln f_t(y_t|w_t,\widehat{\beta}_n)]'$

(ii) Determine a value for $\theta : \widehat{\theta}$

[6] For more motivation see Van der Sluis (1996)

[7] Calculations were performed on a P90 with 16Mb RAM and a P133 with 24 Mb RAM and on some of the 76 RS/6000 nodes of the SP2 computer at SARA (Stichting Academisch Rekencentrum Amsterdam) using programs in Ox (Doornik (1996)) combined with programs in C/C++ written by the author himself. See Van der Sluis (1996 and 1997) for more on practical issues. The programs for the PC and for the RS/6000 will be made available as EmmPack 1.0, see Van der Sluis (1997).

Also PcGive (Hendry and Doornik (1996)) was sometimes used and some analysis was done in Gauss 3.14.

(iii) Simulate $y_1^N(\widehat{\theta})$ and $y_2^N(\widehat{\theta})$ using *antithetic variables*

(iv) Calculate $m_N(\widehat{\theta},\widehat{\beta}_n) = \frac{1}{2N}\sum_{t:=1}^N [\frac{\partial}{\partial\beta}\ln f_t(y_1^N(\widehat{\theta})|w_t,\widehat{\beta}_n) + \frac{\partial}{\partial\beta}\ln f_t(y_1^N(\widehat{\theta})|w_t,\widehat{\beta}_n)]$

(v) Repeat (ii) till (iv) until the quantity $m_N^{'}(\widehat{\theta},\widehat{\beta}_n)(\widehat{\mathcal{I}}_n)^{-1}m_N(\widehat{\theta},\widehat{\beta}_n)$ is minimised

We set $N = 50,000$[8] and as explained in Van der Sluis (1996 and 1997) for this specific stochastic volatility model no starting values have to be used[9]. In this context antithetic variables as a variance reduction technique are also successfully applied in Andersen and Lund (1996). See Ross (1990), among others, for a review of variance reduction techniques. To obtain standard errors for the elements of $\widehat{\theta}_n$ we need to determine $\mathcal{M}(\widehat{\theta}_n,\widehat{\beta}_n)$. This can be done efficiently using trial values. Gallant and Tauchen (1996) advise to fit the *local quadratic regressions*

$$m_i = b_{0i} + b_i^{'}(\theta-\widehat{\theta}_n) + (\theta-\widehat{\theta}_n)^{'}B_i(\theta-\widehat{\theta}_n), i := 1,...,|\beta|.$$

for points near $\widehat{\theta}_n$ and then to take $\mathcal{M}(\widehat{\theta}_n,\widehat{\beta}_n)$ to be the matrix with rows equal to $b_i^{'}$.

Under the null that the structural model is true one may deduce.

$$n\cdot m_N^{'}(\widehat{\theta}_n,\widehat{\beta}_n)(\widehat{\mathcal{I}}_n)^{-1}m_N(\widehat{\theta}_n,\widehat{\beta}_n) \xrightarrow{d} \chi^2_{|\beta|-|\theta|}$$

also the direction of the misspecification may be indicated by the quasi-t ratios

$$\begin{aligned}\widehat{T}_n &:= \widehat{S}_n^{-1}\sqrt{n}m_N(\widehat{\theta}_n,\widehat{\beta}_n)\\ \widehat{S}_n &:= [\text{diag}(\widehat{\mathcal{I}}_n - \widehat{\mathcal{M}}_n(\widehat{\mathcal{M}}_n^{'}\widehat{\mathcal{I}}_n^{-1}\widehat{\mathcal{M}}_n)^{-1}\widehat{\mathcal{M}}_n^{'})]^{1/2}\end{aligned}$$

Here $\widehat{T}_n$ is distributed as $t_{|\beta|-|\theta|}$.

The PSP test is implemented analogously; we have under the null

$$\begin{aligned}&n_2 m_N^{'}(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})\{\widehat{\mathcal{I}}_{n_2} + k\mathcal{M}_2(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})[\mathcal{M}_1^{'}(\widehat{\theta}_{n_1},\widehat{\beta}_{n_1})\widehat{\mathcal{I}}_{n_1}^{-1}\mathcal{M}_1(\widehat{\theta}_{n_1},\widehat{\beta}_{n_1})]^{-1}\mathcal{M}_2^{'}(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})\}^{-1}\cdot\\ &m_N(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2}) \xrightarrow{d} \chi^2_{q_2}\end{aligned}$$

Here we must determine $\mathcal{M}_2$ numerically by simulation since no trial values are available. This may take some time, but only a fraction of the time needed to calculate $\widehat{\theta}_{n_2}$. Note that we have no loss of degrees of freedom since no restrictions are posed on the data to obtain estimates. The direction of misspecification may also be indicated by the quasi t-ratios

$$\widehat{D}_{n_2} := \sqrt{n_2}\text{diag}[\widehat{\mathcal{I}}_{n_2}+k\mathcal{M}_2(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})[\mathcal{M}_1^{'}(\widehat{\theta}_{n_1},\widehat{\beta}_{n_1})\widehat{\mathcal{I}}_{n_1}^{-1}\mathcal{M}_1(\widehat{\theta}_{n_1},\widehat{\beta}_{n_1})]^{-1}\mathcal{M}_2^{'}(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})]^{-1/2}m_N(\widehat{\theta}_{n_1},\widehat{\beta}_{n_2})$$

The $\overleftarrow{\text{PSP}}$ test is implemented in a parallel way to the PSP test.

[8] Two antithetic series are used, each of size 50,000

[9] The model is started up in

$$\begin{aligned}y_0 &= \epsilon_0\sigma_0,\\ \ln\sigma_0^2 &\sim N(\omega/(1-\sum_{i=1}^p\rho_i),\sigma_\eta^2(1+\sum_{j=1}^q\zeta_j^2)(1-\sum_{i=1}^p\rho_i^2)^{-1})\\ \epsilon_0 &\sim N(0,1),\ \text{Corr}(\epsilon_0,\ln\sigma_0^2) = \lambda/\sqrt{(1+\sum_{j=1}^q\zeta_j^2)(1-\sum_{i=1}^p\rho_i^2)^{-1}}\end{aligned}$$

7 Application to exchange rates

The data set under investigation are daily spot prices in Canadian dollar of the British pound, see figure 1. The data are taken from the *Pacific Exchange Rate Server*[10]. The available daily data ranges from January 1971 till August 1996. The interesting breakpoint lies at what is known as Black Wednesday[11] when Britain had to leave the European Monetary System (EMS) and its Exchange Rate Mechanism (ERM). The price-movement series are analysed so we consider $100[\ln(ExchangeRate_t) - \ln(ExchangeRate_{t-1})]$. The post-sample range is set to start at Black Wednesday, consequently the post-sample consists of 1019 daily observations. We decided to perform the PSP test on sample of the same size. Therefore the sample is started at the beginning of September 1988. Some preliminary descriptive statistics can be found in table 1. We observe that both skewness and kurtosis are slightly higher in the post-sample than they are in the sample.

It is reported that exchange rates at a daily rate often do not show autocorrelation in the mean. A correlogram of sample and post-sample, as given in figure 2a and figure 2b, shows that there is some correlation between lags present in the sample and post-sample. However there is no systematic autocorrelation in sample and post-sample except for lag 5 which might represent the *day of the week effect*. The author chooses not to model this effect since his interests are in the volatility process and the effect is weak. This effect will have virtually no impact on the estimation of the volatility process. Moreover simulated series of ARCH or Stochastic Volatility models show the same kind of autocorrelation in the y_t. It would worry us only in case z_t, which are unobserved in the stochastic volatility case, would show this kind of autocorrelation.

The class of auxiliary models is the SNP class of models with an EGARCH leading term. We specified a model using BIC, AIC and HQC. Tables 2, 3, 4 and 5 provide the values of these model selection criteria for the EGARCH(1,1)-H(k_z,0), EGARCH(1,2)-H(k_z,0), EGARCH(1,3)-H(k_z,0) and EGARCH(1,1)-H(k_z,1) auxiliary models, respectively. As can be seen from table 5, convergence was hard for the model with $k_x > 0$, because the likelihood does not monotonically decrease with the number of parameters included. Other models where estimated such as the EGARCH(2,1)-H(k_z,0) and the EGARCH(1,4)-H(k_z,0) but they showed no increase in log likelihood that justifies their use by any standards. For the EGARCH(2,1)-H(k_z,0) it was also observed that it is hard to discern between the two AR parameters γ_1 and γ_2. For the British Pound (BP) sample we find the EGARCH(1,3)-H(4,0) model to be a winner. The parameter estimates for the EGARCH$(1,3)$-H$(4,0)$ model can be found in table 6.

Three different SV models were estimated: SARMAV$(1,0)$, SARMAV$(2,0)$ and SARMAV$(3,0)$. Using the EGARCH$(1,3)$-H$(4,0)$ model we obtain the following estimates for the structural SARMAV$(1,0)$ model:

[10] This is a real goldmine of exchange rates. Many historical exchange rate series are freely available from `http://pacific.commerce.ubc.ca/`

[11] September 16th 1992

$$
\begin{aligned}
y_t &= \sigma_t \epsilon_t \\
\ln \sigma_t^2 &= \underset{(-4.20)}{-.103} + \underset{(34.9)}{.894} \ln \sigma_{t-1}^2 + \underset{(17.4)}{.270} \eta_t
\end{aligned}
$$

Table 8 provides the Hansen J-test for overidentifying restrictions and its individual t-values. The J test has a value of 13.4 with 8 degrees of freedom, we have a P-value of .100 which permits acceptance of this model at a .05 level. Quasi-t ratios for the elements of the score indicate that the elements that belong to α_3,κ_1, a_{20} and a_{40} are somewhat big. This indicates for the α_3 and κ_1 that improvement of the structural model can be made by including more lagged terms in the specification and by introducing some asymmetry in the model. The parameters a_{20} and a_{40} indicate that there is something more to be found in the data. More study of the SNP density may reveal this. Although none of these quasi-t values are significant at a .05 level.

For the SARMAV$(2, 0)$ model we obtained the following estimates

$$
\begin{aligned}
y_t &= \sigma_t \epsilon_t \\
\ln \sigma_t^2 &= \underset{(-9.32)}{-.160} + \underset{(58.5)}{.268} \ln \sigma_{t-1}^2 + \underset{(127)}{.570} \ln \sigma_{t-2}^2 + \underset{(34.4)}{.415} \eta_t
\end{aligned}
$$

In table 8 we find about the same results for the Hansen J test as for the SARMAV$(1, 0)$ model. The quasi-t values are of the parameters of the Hermite polynomial are somewhat closer to zero, indicating that some of the features of the data that where not captured in the SARMAV$(1, 0)$ model, but were captured by the Hermite terms in the auxiliary model, are somewhat more captured by the SARMAV$(2, 0)$ model, but not completely. The J test has a value of 11.6 with 7 degrees of freedom, we have a P value of .113. Since a P value is a monotonic function of the actual evidence against H_0, it is very dangerous to choose the best model of these specifications on basis of the P-values (see Berger and Delampady (1987)). A Likelihood Ratio (LR) test for nested hypotheses is easily deduced from the difference in J tests (see Van der Sluis (1997) for details). This difference has a $\chi^2_{p_0}$ distribution where p_0 denotes the number of parameters in the restricted model. For a test of SARMAV(1,0) against SARMAV(2,0) we obtain LR $= 1.784$, with 3 degrees of freedom we get a P-value of $.618$. Which leads to acceptance of H_0.

For the SARMAV$(3, 0)$ we obtained the following estimates

$$
\begin{aligned}
y_t &= \sigma_t \epsilon_t \\
\ln \sigma_t^2 &= \underset{(-5.84)}{-.094} + \underset{(311)}{1.20} \ln \sigma_{t-1}^2 - \underset{(-335)}{1.14} \ln \sigma_{t-2}^2 + \underset{(108)}{.851} \ln \sigma_{t-3}^2 + \underset{(31.9)}{.267} \eta_t
\end{aligned}
$$

From table 8 we now see that the quasi t ratio corresponding to α_3 is much closer to zero. The quasi $t-$ratios corresponding to the parameters in the Hermite polynomial are also closer to their neutral value. This is reflected in the value of the J test which is now 6.87 with 6 degrees of freedom we obtain a P-value of .333. A LR test of SARMAV(1,0) against SARMAV(3,0) has a value of 6.54 which means with 3 degrees of freedom a P-value of .088, which still means acceptance of H_0 at a .05 level. Note that typically the LR test becomes more stringent with an expansion of the alternative. For completeness, the LR test of SARMAV(2,0) against SARMAV(3,0) has a value of 4.7565, with 4 degrees of freedom, this results in a P-value of .313.

We also used the auxiliary EGARCH$(1, 3)$-H$(4, 0)$ model for the post-sample (1019 observations). This model is not BIC-optimal for the post-sample. This in contrast with Van der Sluis (1996) where the specification of the post-sample was also chosen BIC optimal. In Van der Sluis (1996) however test statistics for structural stability were considered for which an optimality property can be proven in case the post-sample model is modelled according to AIC, BIC or HQC. To the author's knowledge for the PSP test no such property exists. For interpretational reasons the same model was specified for the post-sample as for the sample. The moment conditions that were used for estimation of the sample are now evaluated under the post-sample. However there are no urgent reasons to take either a post-sample model that is optimal for some model selection criterion or to take a post-sample model that equals the sample model. The only thing that changes is the null and alternative hypothesis. In principle any model can be used, the only question that matters is whether there are interesting new insights gained by a rejection of the test based on such moment conditions.

For the PSP test, results are provided in table 9. The PSP-test for the SARMAV$(1, 0)$ $(n_1 = 1019, n_2 = 1019)$ has a value of 48.0 which means that on basis of this PSP-test we reject the null hypothesis of structural stability at any reasonable level. Inspection of the $t-$values of the individual components of the PSP test reveals that the elements of the leading term are causing the rejection, not the elements of the SNP density. We may conclude that the asymmetry which became more prominent in the post-sample than it was in the sample is one of the causes of the rejection. The other cause will be a change of the long run volatility $\omega/(1-\rho)$, which is captured by the parameters α_0 and γ_1 in the EGARCH term. For the SARMAV$(2, 0)$ and SARMAV$(3, 0)$ the conclusions are virtually the same.

8 Conclusion

In this paper the post-sample prediction tests for GMM based methods are applied to EMM based methods. One advantage of the PSP test is that this may not only detect parameter instability but it may also serve as a guide in the determination of the dimension of the auxiliary model in case more data becomes available. That is the PSP test enables us to test whether the moment conditions used for estimation are still valid. Generally, the choice of an in-some-sense-optimal auxiliary model for EMM procedures needs more study. This is not only a question of dimensionality since for non-linear models the concept of dimensionality is blurry. At this moment the best way may be to follow the BIC criterion. Another advantage of the PSP-test is its ability to expand the auxiliary model with the expansion of the post-sample. A more practical advantage of this test is that it is computationally attractive. We have an estimate of $\widehat{\theta}_{n_1}$ and simulations from the last run $\{y_\tau(\widehat{\theta}_{n_1})\}_{t:=1}^{N}$ at our disposal. Only an estimate $\widehat{\beta}_{n_2}$ of the parameters in the auxiliary model for post-sample, the score of the auxiliary model in $\widehat{\beta}_{n_2}$ and $\{y_\tau(\widehat{\theta}_{n_1})\}_{t:=1}^{N}$ and the matrices $\mathcal{M}_1(\widehat{\theta}_{n_1}, \widehat{\beta}_{n_1})$ and $\mathcal{M}_2(\widehat{\theta}_{n_1}, \widehat{\beta}_{n_2})$ must be determined. However this will take only a fraction of the time needed to determine $\widehat{\theta}_{n_2}$. This is even more important for EMM based techniques than it is for GMM based techniques.

The application is made to the British pound versus Canadian dollar exchange rates. The structural brakpoint was set at September 16th 1992 where Britain was forced to leave the European Monetary Union, Exchange Rate Mechanism. The null hypothesis of structural stability was clearly rejected. Recently Britain has returned to the ERM. Therefore it may be interesting to investigate another structural breakpoint when more data becomes available as time progresses. PSP tests may be also be applied to multiple breakpoints, in case the asymptotics provide good approximations for finite sample sizes.

Besides the tests for structural stability this paper gives a successful application of the EMM methodology to several stochastic volatility models.

A standard drawback of the PSP test is the fact that the breakpoint is assumed known. This yoke may be shed following work of Ghysels *et al*. (1995). However with the current state of computing power, computing times will be very high, so it will take some years before this could be applied as a standard test procedure.

A Appendix

For the proof to be correct, *regularity conditions* as those in Ghysels and Hall (1990a) must be met. The proof hinges on the δ-method, this may not be the most subtle method, for other methods see Gallant and White (1988).

Proof of theorem 1. For EMM the rationale goes along the same lines. Under stationarity we have

$$\begin{pmatrix} \sqrt{n_1} m_N(\theta_0, \widehat{\beta}_{n_1}) \\ \sqrt{n_2} m_N(\theta_0, \widehat{\beta}_{n_2}) \end{pmatrix} \xrightarrow{d} N(\begin{pmatrix} 0 \\ \delta_2 \end{pmatrix}, \begin{pmatrix} \mathcal{I}_1 & 0 \\ 0 & \mathcal{I}_2 \end{pmatrix}) \text{ as } n_1 \to \infty,\ k = \frac{n_2}{n_1} < \infty \tag{2}$$

The δ-method yields,

$$\sqrt{n_2} m_N(\widehat{\theta}_{n_1}, \widehat{\beta}_{n_2}) \xrightarrow{d} \sqrt{n_2} m_N(\theta_0, \beta_2) + \sqrt{n_2} \mathcal{M}_2(\widehat{\theta}_{n_1}, \widehat{\beta}_{n_2})(\widehat{\theta}_{n_1}(\mathcal{I}_0) - \theta_0) + o_p(1)$$

so

$$\sqrt{n_2} m_N(\widehat{\theta}_{n_1}, \widehat{\beta}_{n_2}) \xrightarrow{d} \sqrt{n_2} m_N(\theta_0, \beta_2) + \sqrt{k}\sqrt{n_1} \mathcal{M}_2(\widehat{\theta}_{n_1}, \widehat{\beta}_{n_2})(\widehat{\theta}_{n_1}(\mathcal{I}_0) - \theta_0) + o_p(1)$$

and because

$$\sqrt{n_1}(\widehat{\theta}_{n_1}(\mathcal{I}) - \theta_0) \xrightarrow{d} N(0, [\mathcal{M}_1'(\theta_0, \beta_1)\mathcal{I}_1^{-1}\mathcal{M}_1(\theta_1, \beta_1)]^{-1})$$

we obtain using (2)

$$\sqrt{n_2} m_N(\widehat{\theta}_{n_1}, \widehat{\beta}_{n_2}) \xrightarrow{d} N(\delta_2, \mathcal{I}_2 + k\mathcal{M}_2'(\theta_0, \beta_2)[\mathcal{M}_1'(\theta_0, \beta_1)\mathcal{I}_1^{-1}\mathcal{M}_1(\theta_0, \beta_1)]^{-1}\mathcal{M}_2(\theta_0, \beta_2))$$

the noncentrality parameter $\lambda_{\mathsf{PSP}_{\mathsf{EMM}}}$ can be found by applying theorem 1 in Newey (1985). Here $\beta_i = \text{plim}_{n_i \to \infty} \widehat{\beta}_{n_i}$. ■

Proof of theorem 2.

As above with the subscripts n_1 and n_2 reversed. ■

References

[1] Ahn, S. C., (1995), Model specification testing based on root-t consistent estimators. Working paper, Arizona State University.

[2] Akaike, H, (1973), Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, 267-281. Budapest: Akademia Kiado.

[3] Andersen, T. B. and J. Lund (1996), Estimating continuous time stochastic volatility models of the short term interest rate, forthcoming *Journal of Econometrics*.

[4] Andersen, T. B. and B. E. Sørensen (1996), Estimation of a stochastic volatility model: A Monte Carlo Study. *Journal of Business and Economic Statistics* **14**, 328-352.

[5] Andrews, D. W. K. and R. C. Fair (1988), Inference in econometric models with structural change. *Review of Economic Studies* **55**, 615-640.

[6] Berger, J. O. and M. Delampady (1987), Testing precise hypotheses, *Statistical Science* **2**, 317-352.

[7] Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307-327.

[8] Bollerslev, T., R. Y. Chou and K. F. Kroner (1992), ARCH modelling in finance: A review of the theory and empirical evidence. *Journal of Econometrics* **52**, 5-59.

[9] Bollerslev, T., R. F. Engle and D. B. Nelson (1994), ARCH models. Ch. 49 in Handbook of Econometrics volume IV. Amsterdam: North-Holland.

[10] Doornik, J. A. (1996), *Object-Oriented Matrix Programming using Ox*, London: International Thomson Business Press and Oxford: `http://www.nuff.ox.ac.uk/Users/Doornik/`

[11] Engle, R. F. (1982), Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica* **50**, 987-1008.

[12] Fenton, V. M. and A. R. Gallant (1996a), Qualitative and asymptotic performance of SNP density estimators. Working Paper.

[13] Fenton, V. M. and A. R. Gallant (1996b), Convergence rates of SNP density estimators. *Econometrica* **64**, 719-727.

[14] Gallant, A. R. and D. W. Nychka (1987), Semi-nonparametric maximum likelihood estimation, *Econometrica* **55**, 363-390.

[15] Gallant, A. R. and H. White (1988), *A unified theory of estimation and inference for nonlinear dynamic models*. Oxford: Basil Blackwell.

[16] Gallant, A. R. and G. E. Tauchen (1989), Seminonparametric estimation of conditionally constrained heterogeneous processes: asset pricing applications, *Econometrica* **57**, 1091-1120.

[17] Gallant. A. R., D. A. Hsieh and G. E. Tauchen (1991), On fitting a recalcitrant series: the pound/dollar exchange rate, 1974-1983, in W.A. Barnett, J. Powell and G. E. Tauchen (eds), *Nonparametric and semiparametric methods in econometrics and statistics*, Fifth International Symposium in Economic Theory and Econometrics, 199-240, Cambridge: Cambridge University Press.

[18] Gallant. A. R., D. A. Hsieh and G. E. Tauchen (1994), Estimation of stochastic volatility models with diagnostics. Forthcoming in *Journal of Econometrics.*

[19] Gallant, A. R. and G. E. Tauchen (1996), Which moments to match? Working paper. Duke University. *Econometric Theory*, in press.

[20] Gallant, A. R. and J. Long (1996), Estimation of continuous time models for stock returns and interest rates. Working paper, University of North Carolina.

[21] Ghysels, E., A. Guay and A. Hall (1995), Predictive tests for structural change with unknown breakpoint. Working paper.

[22] Ghysels, E., A. Harvey and E. Renault (1996), Stochastic Volatility. Forthcoming in *Handbook of Statistics*, vol 14: Statistical Methods in Finance.

[23] Ghysels, E. and A. Hall (1990a), A test for structural stability of Euler conditions parameters estimated via the generalized method of moments estimator. *International Economic Review*, **31**, 355-364.

[24] Ghysels, E. and A. Hall (1990b), Are consumption-based intertemporal capital asset pricing models structural? *Journal of Econometrics* **45**, 121-139.

[25] Gourieroux, C., A. Monfort and E. Renault (1993), Indirect Inference, *Journal of Applied Econometrics* **8**, S85-S119.

[26] Hall, A. and A. Sen (1996), Structural stability testing in models estimated by generalized method of moments. Working paper, North Carolina State University.

[27] Hannan, E. J. and B. G. Quinn, The determination of the order of an autoregression, *Journal of the Royal Statistical Society*, **B 41**, 190-195.

[28] Hansen, L. P. (1982), Large sample properties of generalized method of moments estimators. *Econometrica* **50**, 1029-1054

[29] Harvey, A., E. Ruiz and N. Shephard (1994), Multivariate stochastic models. *Review of Economic Studies*, **61**, 247-264.

[30] Hendry, D. F. and J. A. Doornik (1996), *Empirical econometric modelling using PcGive for Windows*. London: International Thomson Business Press.

[31] Hoffman, D. and A. Pagan (1989), Post-sample prediction tests for generalized method of moments estimators. *Oxford Bulletin of Economics and Statistics*, **51**, 331-43

[32] Kleibergen, F. and H. K. van Dijk (1993), Non-stationarity in GARCH models: a Bayesian analysis, *Journal of Applied Econometrics* **8**, S41-S61.

[33] Lütkepohl, H. *Introduction to multiple time series analysis*. Berlin: Springer Verlag

[34] Nelson D. B., and C. Q. Cao (1992), Inequality constraints in the univariate GARCH model. *Journal of Business & Economics Statistics* **10**, 229-235

[35] Newey, W. K. (1985), Generalized method of moments specification testing, *Journal of Econometrics* **29**, 229-256.

[36] Phillips, P. C. B. (1983), ERA's: A new approach to small sample theory, *Econometrica* **51**, 1505-1527.

[37] Quinn, B. G. (1980), Order determination for a multivariate autoregression, *Journal of the Royal Statistical Society*, **B 42**, 182-185.

[38] Ross, S. M. (1990), *A course in simulation*, New York: MacMillan.

[39] Schwarz, G. (1978), Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.

[40] Sowell, F. (1996), Test for violations of moment conditions. Manuscript, Graduate School of Industrial Administration, Carnegie Mellon University.

[41] Tauchen, G. E. (1996), New minimum-chi-square methods in empirical finance, invited paper at Seventh World Congress of the Econometric Society, Tokyo Japan, August 1995.

[42] Van der Sluis, P. J. (1996), Computationally attractive stability tests for the efficient method of moments, TI Discussion paper no. 96-?. Paper presented at EC^2 on *Simulation Methods in Econometrics*, Florence, Italy, December 12-14 1996.

[43] Van der Sluis, P. J. (1997), EmmPack 1.0: C code for use with Ox for estimation of univariate stochastic volatility models with the efficient method of moments. Unpublished working paper and freeware, University of Amsterdam.

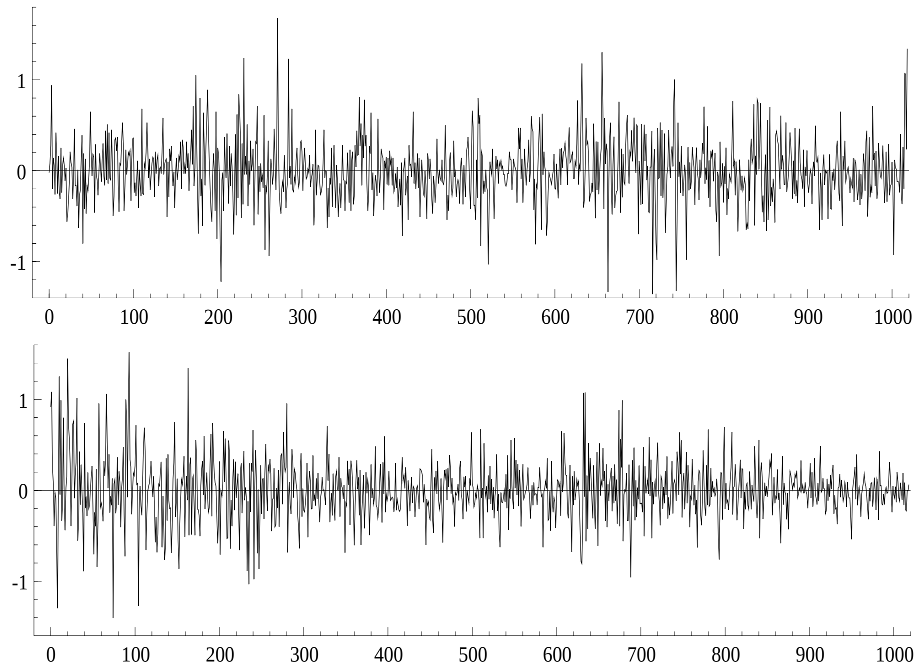

Figure 1: Sample and post-sample

	sample	post-sample
Mean	-.005	.003
Std.Dv.	.695	.664
Skewness	.269	.316
Excess Kurtosis	1.81	2.49
Minimum	-2.61	-2.83
Maximum	3.12	3.04
Normality Chiˆ2	85.7	137

Table 1: Some preliminary statistics of the sample and of the post-sample

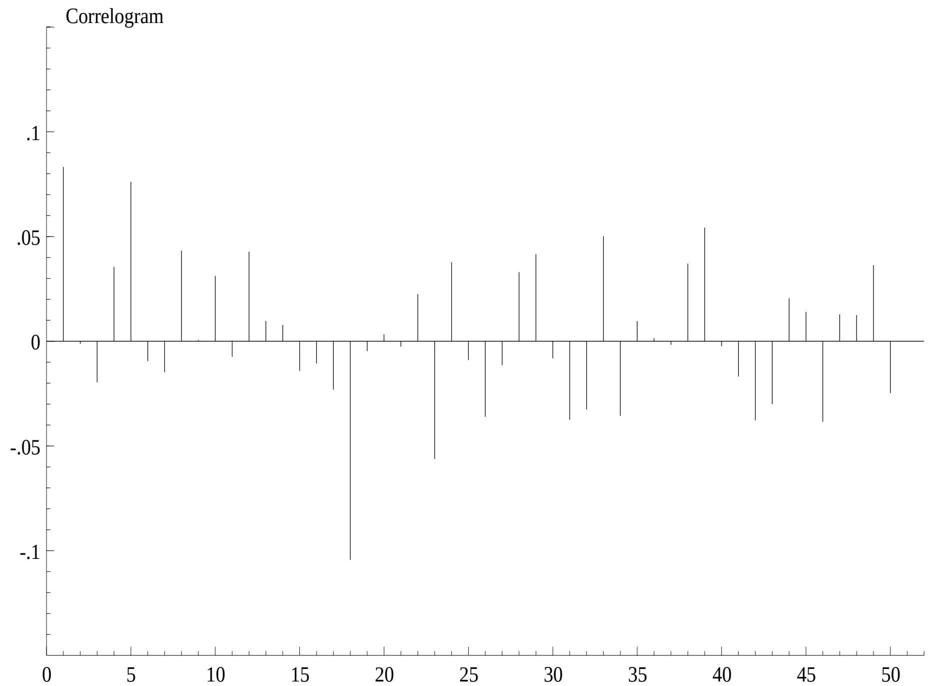

Figure 2a: Correlogram of the sample

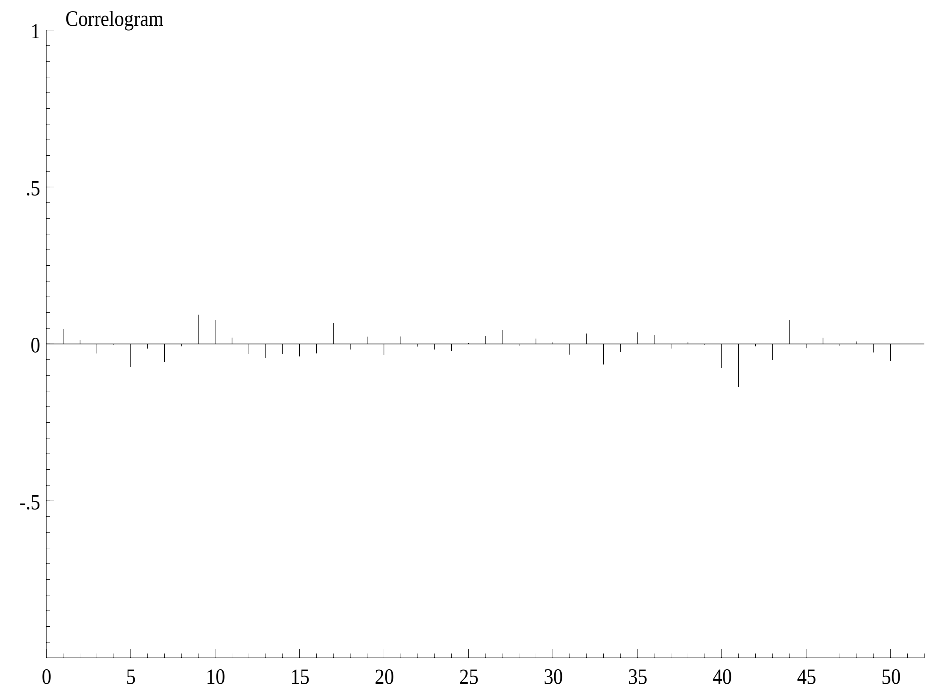

Figure 2a: Correlogram of the post-sample

K_z	Log Likl	AIC	HQC	BIC
1	-1049.2	-1055.2	-1060.8	-1070.0
2	-1049.2	-1056.2	-1062.7	-1073.4
3	-1047.5	-1055.5	-1063.0	-1075.2
4	-1026.5	-1035.5	-1044.0	-1057.7
5	-1026.5	-1036.5	-1045.8	-1061.1
6	-1026.5	-1037.5	-1047.8	-1064.6
7	-1024.5	-1036.5	-1047.7	-1066.1
8	-1024.1	-1037.1	-1049.3	-1069.1
9	-1023.5	-1037.5	-1050.6	-1071.9
10	-1023.5	-1038.5	-1052.5	-1075.4

Table 2: Various model selection criteria for the EGARCH$(1,1)$-H$(K_x, 0)$ model for the sample

K_z	Log Likl	AIC	HQC	BIC
1	-1046.2	-1053.2	-1059.8	-1070.4
2	-1046.2	-1054.2	-1061.7	-1073.9
3	-1044.6	-1053.6	-1062.0	-1075.7
4	-1024.0	-1034.0	-1043.3	-1058.6
5	-1024.0	-1035.0	-1045.3	-1062.1
6	-1023.9	-1035.9	-1047.2	-1065.5
7	-1021.9	-1034.9	-1047.1	-1066.9
8	-1021.2	-1035.2	-1048.3	-1069.7
9	-1020.0	-1035.0	-1049.0	-1071.9
10	-1017.9	-1033.9	-1048.9	-1073.3

Table 3: Various model selection criteria for the EGARCH$(1,2)$-H$(K_z, 0)$ model for the sample

K_z	Log Likl	AIC	HQC	BIC
1	-1039.6	-1047.6	-1055.1	-1067.3
2	-1039.6	-1048.6	-1057.0	-1070.8
3	-1038.2	-1048.2	-1057.5	-1072.8
4	-1018.7	-1029.7	-1039.9	-1056.7
5	-1018.7	-1030.7	-1041.9	-1060.2
6	-1018.7	-1031.7	-1043.8	-1063.7
7	-1016.9	-1030.9	-1044.0	-1065.4
8	-1016.4	-1031.4	-1045.4	-1068.3
9	-1015.6	-1031.6	-1046.6	-1071.0
10	-1015.2	-1032.2	-1048.0	-1074.0

Table 4: Various model selection criteria for the EGARCH$(1,3)$-H$(K_z, 0)$ model for the sample

K_z	Log Likl	AIC	HQC	BIC
1	-1045.0	-1053.0	-1060.4	-1072.7
2	-1044.8	-1054.8	-1064.2	-1079.4
3	-1042.7	-1054.7	-1065.9	-1084.2
4	-1021.9	-1035.9	-1049.0	-1070.4
5	-1028.5	-1044.5	-1059.5	-1083.9
6	-1024.6	-1042.6	-1059.4	-1086.9
7	-1019.7	-1039.7	-1058.4	-1089.0
8	-1030.7	-1052.7	-1073.3	-1106.9
9	-1025.2	-1049.2	-1071.6	-1108.3
10	-1018.4	-1044.4	-1068.7	-1108.4

Table 5: Various model selection criteria for the EGARCH$(1,1)$-H$(K_z,1)$ model for the sample

	parameters	t-values
α_0	-.088	-8.72
α_1	-.550	-2.50
α_2	.157	.635
α_3	.827	2.95
γ_1	.883	77.6
κ_1	.042	1.90
κ_2	.158	4.29
a_{10}	-.015	-.949
a_{20}	.006	-.414
a_{30}	.018	1.09
a_{40}	.089	5.64

Table 6: Sample estimates for the parameters of the EGARCH$(1,3)$-H$(4,0)$ model.

	parameters	t-values
α_0	-.010	-6.82
α_1	-.568	-7.70
α_2	.409	5.35
α_3	-.272	3.45
γ_1	1.00	1012
κ_1	-.050	-2.98
κ_2	.102	6.85
a_{10}	-.016	-.977
a_{20}	.101	6.30
a_{30}	.007	.457
a_{40}	.126	8.15

Table 7: Post-sample estimates for the parameters of the EGARCH$(1,3)$-H$(4,0)$ model.

	SARMAV(1,0)	SARMAV(2,0)	SARMAV(3,0)
	J-test	J-test	J-test
X	13.4	11.6	6.87
df	8	7	6
$\Pr(X > \chi^2)$	.100	.113	.333
α_0	-.846	.847	-.259
α_1	.397	.435	-1.43
α_2	-.754	-.950	-1.17
α_3	-2.28	-2.22	-1.31
γ_1	-.271	.258	-.333
κ_1	-1.79	-1.78	-1.95
κ_2	.030	-.755	-.220
a_{10}	.951	.942	.914
a_{20}	-1.71	-1.38	-.649
a_{30}	-1.17	-1.14	-1.25
a_{40}	-1.84	-1.37	-.433

Table 8: Individual t values of the elements of the J statistic. We have $\Pr(|t_6| \geq 2.447) = 0.05$, $\Pr(|t_7| \geq 2.365) = 0.05$ and $\Pr(|t_8| \geq 2.306) = 0.05$.

	SARMAV(1,0)	SARMAV(2,0)	SARMAV(3,0)
	PSP-test	PSP-test	PSP-test
X	48.0	46.3	42.0
df	11	11	11
$\Pr(X > \chi^2)$	.000	.000	.000
α_0	2.54	2.72	2.94
α_1	1.90	2.10	2.25
α_2	1.72	1.94	2.32
α_3	1.72	1.96	2.41
γ_1	-2.86	-3.01	-3.20
κ_1	3.40	3.58	3.59
κ_2	3.08	3.17	3.36
a_{10}	1.02	.997	1.01
a_{20}	1.85	1.66	1.47
a_{30}	-.568	-.568	-.606
a_{40}	-1.53	-.995	-.551

Table 9: Individual t values of the elements of the PSP statistic. We have $\Pr(|\mathrm{t}_{11}| \geq 2.201) = 0.05$