UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 39

Essays on Time Series Analysis

With Applications to Financial Econometrics

DANIEL PREVE

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2008

ISSN 1652-9030
ISBN 978-91-554-7158-3
urn:nbn:se:uu:diva-8638

Dissertation presented at Uppsala University to be publicly examined in Hörsal 2, Ekonomikum, Kyrkogårdsgatan 10, Uppsala, Friday, May 9, 2008 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract
Preve, D. 2008. Essays on Time Series Analysis. With Applications to Financial Econometrics. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 39. 21 pp. Uppsala.
ISBN 978-91-554-7158-3.

This doctoral thesis is comprised of four papers that all relate to the subject of Time Series Analysis.

The first paper of the thesis considers point estimation in a nonnegative, hence non-Gaussian, AR(1) model. The parameter estimation is carried out using a type of extreme value estimators (EVEs). A novel estimation method based on the EVEs is presented. The theoretical analysis is complemented with Monte Carlo simulation results and the paper is concluded by an empirical example.

The second paper extends the model of the first paper of the thesis and considers semiparametric, robust point estimation in a nonlinear nonnegative autoregression. The nonnegative AR(1) model of the first paper is extended in three important ways: First, we allow the errors to be serially correlated. Second, we allow for heteroskedasticity of unknown form. Third, we allow for a multi-variable mapping of previous observations. Once more, the EVEs used for parameter estimation are shown to be strongly consistent under very general conditions. The theoretical analysis is complemented with extensive Monte Carlo simulation studies that illustrate the asymptotic theory and indicate reasonable small sample properties of the proposed estimators.

In the third paper we construct a simple nonnegative time series model for realized volatility, use the results of the second paper to estimate the proposed model on S&P 500 monthly realized volatilities, and then use the estimated model to make one-month-ahead forecasts. The out-of-sample performance of the proposed model is evaluated against a number of standard models. Various tests and accuracy measures are utilized to evaluate the forecast performances. It is found that forecasts from the nonnegative model perform exceptionally well under the mean absolute error and the mean absolute percentage error forecast accuracy measures.

In the fourth and last paper of the thesis we construct a multivariate extension of the popular Diebold-Mariano test. Under the null hypothesis of equal predictive accuracy of three or more forecasting models, the proposed test statistic has an asymptotic Chi-squared distribution. To explore whether the behavior of the test in moderate-sized samples can be improved, we also provide a finite-sample correction. A small-scale Monte Carlo study indicates that the proposed test has reasonable size properties in large samples and that it benefits noticeably from the finite-sample correction, even in quite large samples. The paper is concluded by an empirical example that illustrates the practical use of the two tests.

Keywords: non-Gaussian time series, nonnegative autoregression, robust estimation, strong convergence, realized volatility, volatility forecast, forecast comparison, Diebold-Mariano test

Daniel Preve, Department of Information Science, Box 513, Uppsala University, SE-75120 Uppsala, Sweden

ISSN 1652-9030
ISBN 978-91-554-7158-3
urn:nbn:se:uu:diva-8638 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-8638)

To My Family.

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Preve, D. (2007) Point Estimation in a Nonnegative First-Order Autoregression.

II Preve, D. (2007) Robust Point Estimation in a Nonnegative Autoregression.

III Eriksson, A., Preve, D., Yu, J. (2007) Forecasting Realized Volatility using a Nonnegative Semiparametric Model.

IV Mariano R. S., Preve, D. (2007) Statistical Tests for Multiple Forecast Comparison.

Contents

1. Introduction

This doctoral thesis is comprised of four papers that all relate to the subject of Time Series Analysis.

A time series is a set of observations ordered by time. In the very simplest case, a time series is a sequence of recorded values of one variable taken at equally spaced time points. For example, the (time ordered) sequence of daily closing prices of the Apple Inc. stock is a time series. Time series can be found in the fields of engineering, science, sociology and economics.

Time series analysis is a branch of statistics which deals with techniques developed for drawing inferences from time series. The first step in the analysis of a time series is the selection of a suitable model (or class of models) for the data. To allow for the unpredictable nature of future observations it is assumed that each observation is a realized value of a random variable. Given a particular time series, the primary goals of time series analysis are: (1) to set up a hypothetical statistical model to represent the series in order to obtain insights into the mechanism that generates the data[1] and (2), once a satisfactory model has been formulated, to extrapolate from the model in order to anticipate (forecast) the future values of the time series.

For example, a time series econometrician faces the task to construct models capable of forecasting, interpreting, and testing hypothesis concerning economic data.[2]

Having selected a time series model the parameters of the model need to be estimated and its goodness of fit to the data has to be checked. The first two papers of this thesis concerns parameter estimation in a class of nonnegative time series models. If the model is satisfactory it may be used for forecasting. In the third paper of this thesis we construct a simple nonnegative model for certain financial time series data, use the results of the second paper to estimate the proposed model on empirical data, and then use the estimated model to make forecasts. Once a time series has been analyzed and its future values have been forecasted, it is reasonable to question how good the forecasts are. Typically, there will be several plausible models to extrapolate from in order to forecast the series. The fourth and last paper of this thesis constructs a test for multiple forecast comparison.

[1]However, whether the real life process generating the data can be reliably and completely represented in terms of a statistical model is a different matter altogether. It has been argued that there never is an attainable true data generating process and that the best that can be hoped for is that a very restricted class of models can be successfully used.

[2]Depending on the particular field of application, other applications include separation of noise from signals and the control of future values of a series.

2. Summary of Papers

2.1 Paper I

Setting

A time series model is a natural model for describing real life processes and their time series. One particular class of time series models plays a central part in this thesis: autoregressive models. Let X_t denote the value of a data point at period t. The simplest example of an autoregression is the first-order autoregressive, abbreviated AR(1), model given by the relation

$$X_t = \phi X_{t-1} + Z_t, \tag{2.1}$$

for $t = 2, 3, 4$ and so on. In this model, ϕ (the autoregressive parameter) controls the persistence in the model and the Z_t (the 'errors') are random variables assumed to be mutually independent, identically distributed and independent of X_1 (the initial value). Traditionally $Z_2, Z_3, \ldots$ are assumed to be Gaussian distributed with mean zero.

For example, (2.1) has been proposed as a model for daily stock prices. The autoregressive parameter ϕ is then usually assumed to be 1 reflecting that, in an efficient market, the best forecast of tomorrows stock price is the current price (day-to-day changes in the price of a stock should have an expected value of zero). This model is known as the Random Walk model.

The AR(1) model is often further adjusted to accommodate for trend(s) in the data by the addition of a dynamic trend component μ_t, which allows for a long-term change in the mean level of the process. The model then becomes

$$X_t = \mu_t + \phi X_{t-1} + Z_t.$$

The addition of a trend component needs to be further motivated. Typically, a time series is considered to be composed of four types of components: the trend, the cycle, the seasonal variation (for sub annual data) and an irregular component.[1] The trend is generally thought of as a smooth and slow movement over a long term (for instance, there is empirical evidence that even though stock prices move up and down randomly there is over time, however, an upward trend). The addition of a trend component can improve the fit and forecast accuracy of the model (because any predictable component can be extrapolated into the future).

[1] In our model the irregular component is $\phi X_{t-1} + Z_t$ and the seasonal and cyclical components are zero.

Due to the recursive nature of X_t, an alternative representation is given in terms of the initial value X_1 by

$$X_t = \phi^{t-1} X_1 + \sum_{k=0}^{t-2} \phi^k (\mu_{t-k} + Z_{t-k}).$$

If X_1 is fixed (or Gaussian distributed), the trend component is deterministic, and the errors are Gaussian, then X_t too is Gaussian. However, a Gaussian model may assume negative values, which is not a very desirable property of a price, duration or a volatility.

If it is known that the values $X_1, X_2, \ldots$ must be nonnegative (hence non-Gaussian), then the following, restricted, AR(1) specification can be used

$$\begin{cases} X_t = \mu_t + \phi X_{t-1} + Z_t, \\ \mu_t \geq 0 \text{ for all } t, \\ \phi \geq 0, \\ X_1 > 0 \text{ with probability } 1, \\ Z_2, Z_3, \ldots \text{ are nonnegative.} \end{cases}$$

Well-known examples of nonnegative random variables include exponential, lognormal and inverse Gaussian random variables.

Contribution

Autoregressive moving average (ARMA) model building is usually carried out under the assumption that the time series observations are Gaussian distributed, even though the use of a Gaussian error distribution does not adjust the distribution of the ARMA model to account for non-Gaussianity in the data generating process. Consequently, nonnegative time series data is usually transformed in order to make it appear Gaussian distributed. See, for example, [2] and [3]. This approach would typically involve the estimation of one or more transformation parameters, resulting in a nonlinear model specification. Because any inference based on a transformation from $(0, \infty)$ to $(-\infty, \infty)$ potentially ignores the nonnegative nature of the original observations, it could be argued that this approach does not always take all available information into account (cf. Figures 2.2 and 2.3). In contrast, nonnegative ARMA models have the potential to model nonnegative observations directly and more parsimoniously.

Given the sample $X_1, \ldots, X_T$ we are interested in the selection and estimation of a suitable model (or class of models) for the data. The first paper of this thesis considers point estimation in the nonnegative AR(1) model. It is shown that the extreme value estimator (EVE) $\min\{X_t/X_{t-1}\}_{t=2}^{T}$ of the autoregressive coefficient, suggested in [6] and [8] among others, is robust in the presence of an unknown time-varying trend component. Two natural extensions of the EVE are also proposed, for the exceptional situation when the nonnegative support of the error is known and different from $[0,\infty)$, and sufficient conditions for the EVEs to coincide with the Maximum Likelihood (ML) counterpart are given. It is noted that the derivation of ML estimators for the nonnegative AR(1) model generally is analytically infeasible. In recognition of this inconvenience a novel estimation method, the Perturbed Maximum Likelihood (PML) method, is presented. The theoretical analysis of the paper is complemented with Monte Carlo simulation results. Simulation studies illustrate the asymptotic theory and indicate reasonable small sample properties of the proposed estimators. The paper is concluded by an empirical example that illustrates a PML based inference procedure.

We remark that, in view of the robustness result of the paper, it follows that the strong consistency[2] of the EVE holds also when an unknown nonnegative deterministic seasonal (or cyclical) component is added to the model specification.

2.2 Paper II

Setting

In the last two decades, nonlinear and also nonstationary times series models have gained much attention. This interest is mainly motivated by the fact that there is empirical evidence that many real life time series are non-Gaussian and have a structure that change over time.[3] For example, many economic time series are known to show a large number of nonlinear features such as cycles, asymmetries, jumps, thresholds, heteroskedasticity[4] and combinations thereof, that additionally need to be taken into account.

Contribution

The second paper of this thesis extends the model in Paper I and considers semiparametric, robust estimation in a nonlinear nonnegative autoregression, that may be a useful tool in describing the behavior of a broad class of nonnegative time series. In some applications, robust estimation of the autoregressive coefficient ϕ is of interest in its own right. One example is point forecasting, as described in Paper III of this thesis. In recognition of this fact, Paper II focuses explicitly on the consistent and robust estimation of ϕ. In this paper, we extend the nonnegative AR(1) model of Paper I in three important ways: First, we allow the errors to be m-dependent of unknown order m (successive errors no longer have to be stochastically independent). The property of m-dependence generalizes that of independence in a natural way. Observations of

[2]The estimator $\hat{\theta}_T$ of the parameter θ is said to be strongly consistent if $Pr(\lim_{T\to\infty}\hat{\theta}_T=\theta)=1$.
[3]Since, for example, an economy changes due to unforeseen interventions, it is difficult to justify using the same model over a longer period of time.
[4]A sequence of random variables is heteroskedastic if the random variables have different variances (the complementary concept is called homoskedasticity).

an m-dependent process are independent provided they are separated in time by more than m time units.[5] This is important as the actual dynamics of a time series is often more complex than the dynamics of an AR(1), and the original model can be seriously misspecified. Second, we allow for heteroskedasticity of unknown form. This is important since time-varying second moments is a characteristic shared by many different types of time series. Third, we allow for a multi-variable mapping of previous observations. This makes various lagged/seasonal nonlinear model specifications possible.

It is interesting to note that the main result of Paper II can be extended to more general situations. First, in view of the robustness result of Paper I, it should come as no surprise that the modified EVE of Paper II remains strongly consistent if a suitable trend, seasonal or cyclical component (or combinations thereof) is added to the model. Second, it can be shown that the modified EVE is consistent in the presence of certain types of m_t-dependent errors (here the order of the dependence is time-varying). This can be important since it is often difficult to justify using the same model over a longer period of time.

2.3 Paper III

Setting

One task facing the modern time series econometrician is to construct reasonably simple models capable of describing and forecasting economic data. Since financial variables such as stock prices, price durations and volatilities are all inherently nonnegative it is interesting to investigate how well nonnegative time series models are capable of describing financial time series data.

Figure 2.1 plots the monthly realized volatilities[6] (RV) of Standard & Poor's 500 index[7] for the period January 1946 to December 2004. Figures 2.2 and 2.3 shows histograms of the RV and of the logarithmic RV (log-RV). For RV, the departure from Gaussianity is apparent. By contrast, the distribution of log-RV appears to be closer to a Gaussian distribution.

[5] A sequence $U_1, \ldots, U_T$ of random variables is said to be m-dependent if and only if U_t and U_{t+k} are pairwise independent for all $k > m$. In the special case when $m = 0$, m-dependence reduces to independence.

[6] Realized volatility is a measure of the latent historical volatility of a financial instrument, such as a stock or an index. For example, one could calculate the realized volatility for the Apple Inc. stock in Jan of 2008 by taking the standard deviation of its daily returns within that month.

[7] Standard & Poor's 500 index (S&P 500) is an index of 500 stocks chosen for market size, liquidity and industry grouping, among other factors. The S&P 500 is designed to be a leading indicator of U.S. market equities and it is one of the most commonly used benchmarks for the overall U.S. stock market.

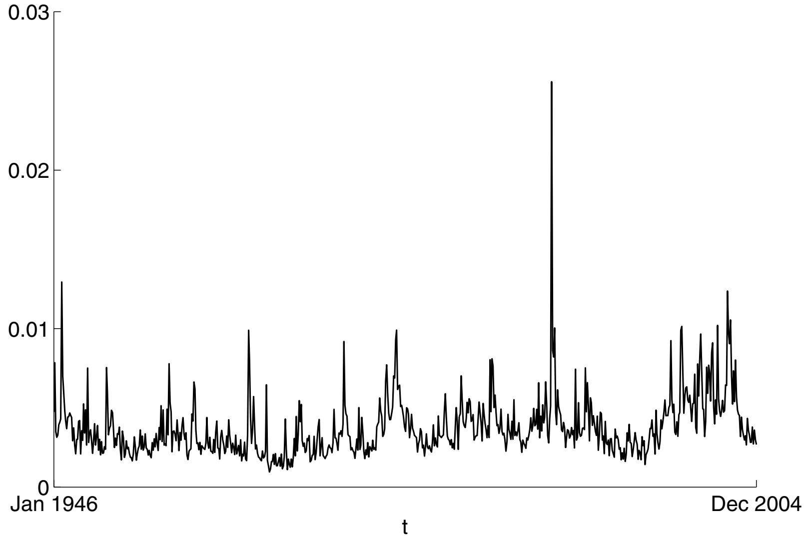

Figure 2.1: S&P 500 monthly realized volatilities, Jan 1946-Dec 2004.

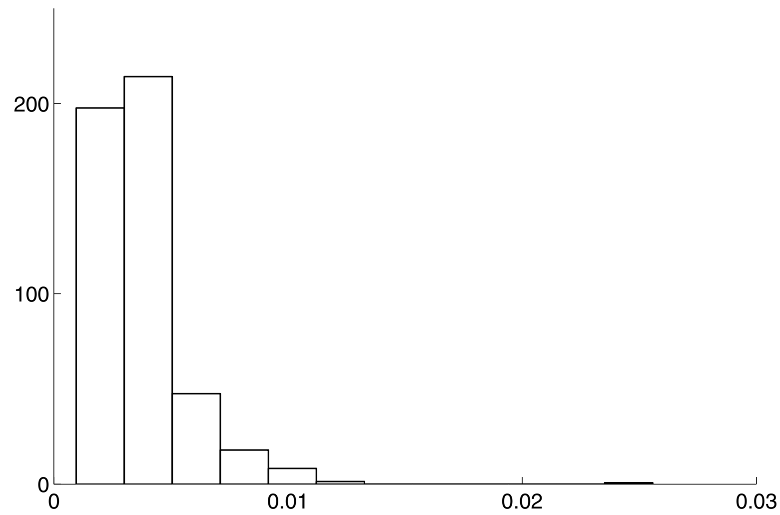

Figure 2.2: Histogram of S&P 500 monthly RV.

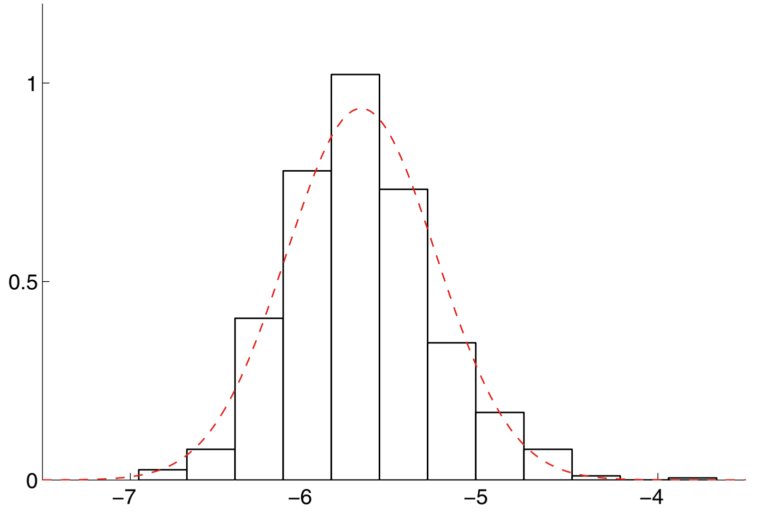

Figure 2.3: Normalized histogram of S&P 500 monthly log-RV with superimposed, estimated Gaussian density curve.

Contribution

In the third paper of this thesis we construct a simple nonnegative model for realized volatility, use the results of Paper II to estimate the proposed model on S&P 500 monthly realized volatilities, and then use the estimated model to make one-month-ahead forecasts.[8] The out-of-sample performance of the proposed model is evaluated against a number of standard models. Various tests and accuracy measures are utilized to evaluate the forecast performances. It is found that forecasts from the new model perform exceptionally well under the mean absolute error and the mean absolute percentage error forecast accuracy measures.

The proposed nonnegative model is of the form

$$\begin{cases} RV_t^{\lambda} = \phi RV_{t-1}^{\lambda} + V_t, \\ \lambda \neq 0, \\ \phi > 0, \\ RV_1 > 0 \text{ with probability } 1, \\ V_2, V_3, \ldots \text{ are nonnegative.} \end{cases}$$

$V_2, \ldots, V_T$ is assumed to be a sequence of m-dependent, identically distributed, continuous random variables with nonnegative support $[\gamma, \infty)$, for some unknown $\gamma \geq 0$ (an intercept in the model is superfluous because γ can be strictly positive). It is assumed that m is finite and potentially unknown. In general,

[8] Returns of stocks are generally thought of as difficult, if not impossible, to predict. In contrast, there is evidence that the volatilities of the returns are relatively easier to forecast.

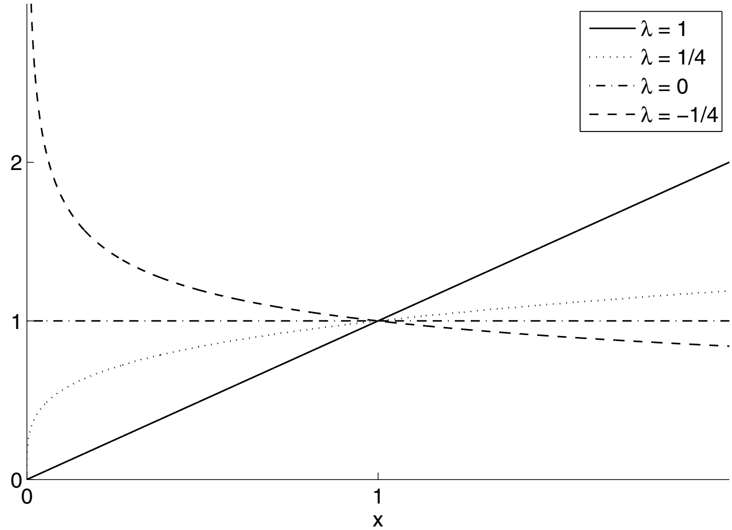

Figure 2.4: Power transform curves for four different values of λ. By the identity $x^{\lambda} = e^{\lambda \ln x}$ for $x > 0$, it is readily seen that x^{λ} is strictly decreasing if $\lambda < 0$ and strictly increasing if $\lambda > 0$. Hence, the power transformation is one-to-one and onto for $x > 0$ and $\lambda \neq 0$.

neither the dependency structure nor the distributional form is assumed to be known for the error V_t. Hence, the model combines a parametric component for the persistence with a nonparametric component for the error. On the one hand, the proposed model is highly parsimonious. In particular, there are only two parameters that have to be estimated for the purpose of volatility forecasting, namely ϕ and λ. On the other hand, the specification is sufficiently flexible for modeling the error. For example, the error is not required to have finite higher order moments and can easily incorporate jumps.

Typically, an MA(m) structure may be assumed for V_t. The presence of a moving average structure of unknown order in the model can be motivated in various ways. For example, [4] showed that volatility can have both persistent and non-persistent components. For another example, effects of various market microstructure noises may not be negligible for estimating RV ([9] and [1]).

One role that the transformation parameter, λ, plays in the proposed model is to stabilize the variance, i.e. to induce homoskedasticity. Figure 2.4 illustrates the power transform function x^{λ} for four different values of λ. For $\lambda < 1$ the transformation tends to suppress larger fluctuations that occur over portions of the time series where the underlying values are larger and may be useful to 'equalize' the variability over the length of a single time series and to improve linearity in the data. However, in contrast to the Box-Cox transformation (for which the logarithmic transformation is a special case) the power transformation maps $(0,\infty)$ to $(0,\infty)$, thus ensuring that the transformed data remains nonnegative.

The transformation parameter controls the nonlinear dependency structure of the model, and allows the conditional variance of RV_t to be time-varying. To see this, suppose that λ is rational.[9] For ease of exposition, suppose that $\lambda = 1/n$ for some natural number n, then

$$\sqrt[n]{RV_t} = \phi \sqrt[n]{RV_{t-1}} + V_t,$$

or equivalently

$$RV_t = (\phi \sqrt[n]{RV_{t-1}} + V_t)^n = \sum_{k=0}^{n} \binom{n}{k} \phi^{n-k} RV_{t-1}^{(n-k)/n} V_t^k,$$

where it appears that the conditional variance of the model depends on the previous realization.

2.4 Paper IV

Setting

Once a time series has been analyzed and its future values have been forecasted, it is reasonable to question how good the forecasts are. Typically, there will be several plausible models to extrapolate from in order to forecast the series. With forecasts from several models it is inevitable that the sample will show differences in forecast accuracy between the different models. Because of this it is important to investigate how likely this outcome is due to pure chance, that is, whether the observed difference is statistically significant or not. If there are just two plausible models, one way to do this is to put the alternative models to a head-to-head test. Since the future values of the time series are unknown, it is reasonable to hold back a portion of the observations from the estimation process and estimate the alternative models over the shortened span of data. These estimates can then be used to forecast the observations of the holdback period, and the properties of the forecast errors of the two models can then be compared.

For example, suppose that an analyst is unsure whether his two alternative models forecast the time series $X_1, \ldots, X_{100}$ equally well or not. One way for the analyst to proceed is to use the first 50 observations to estimate both models and then use the estimates to forecast the value of X_{51}. Since the actual value of X_{51} is known, he can then calculate the forecast error of each model. Next, he can re-estimate the two models using the first 51 observations (this will generally change the parameter estimates obtained in the previous step somewhat) in order to forecast the value of X_{52}. Since the value of X_{52} also is known, he can then calculate two more forecast errors. This scheme can

[9] Recall that any real number can be approximated arbitrarily well by a rational number.

then be continued in order to obtain two distinct time series of one-step-ahead forecast errors, each composed of 50 observations.[10] The analyst can then, for instance, calculate and compare the mean square prediction errors (MSPE) of the two series.

Several tests have been proposed to determine whether the MSPE of one model is statistically different from some other model. In an important contribution, [5] used standard results to derive a test statistic in a more general setting that allows for other measures of forecast accuracy than the MSPE.[11] In their approach, they consider two time series of forecast errors ($e_{i1},...,e_{iT}$ and $e_{j1},...,e_{jT}$ say) and propose a simple test to assess the expected loss associated with each of the forecast series. The quality of each forecast is evaluated by some loss function g of the forecast error. In this setting, the null hypothesis of equal predictive accuracy is $E\,d_t = 0$ where $d_t = g(e_{it}) - g(e_{jt})$. Under fairly weak conditions, they conclude that the test statistic

$$\frac{\bar{d}}{\sqrt{\hat{\omega}/T}},$$

is asymptotically standard Gaussian distributed under the null hypothesis, where $\bar{d}$ is the sample mean of the series $d_1,...,d_T$ and $\hat{\omega}$ is a consistent estimator of the asymptotic variance of $\sqrt{T}\bar{d}$.

Contribution

In the fourth and last paper of this thesis we construct a multivariate extension of the Diebold-Mariano test. Under the null hypothesis of equal predictive accuracy of three or more forecasting models, the proposed test statistic has an asymptotic Chi-squared distribution. To explore whether the behavior of the test in moderate-sized samples can be improved, we also provide a finite-sample correction which simplifies to the finite-sample correction of the Diebold-Mariano test by [7] in the bivariate case. It is pointed out that the correction of Harvey et al. can be further improved. A small-scale Monte Carlo study indicates that the proposed test has reasonable size properties in large samples and that it benefits noticeably from the finite-sample correction, even in quite large samples. The paper is concluded by an empirical example that illustrates the practical use of the two tests.

[10] The scheme described in the text is known as the *recursive* scheme. In the forecasting literature, three schemes for how to generate the sequence of model estimates stand out. The other two are the *rolling* scheme and the *fixed* scheme.

[11] An applied econometrician might be interested in measures of forecast accuracy other than the sum of squared forecast errors. For example, if the loss from making an incorrect forecast depends on the size of the forecast error, it is more natural to consider the absolute values of the forecast errors (using the squared errors makes sense only if the loss is quadratic).

3. Acknowledgements

Mentioned here is only a subset of the people I owe thanks to. Firstly, I would like to thank my colleagues at the Department of Information Science for many fun times and great memories. A special thanks to my thesis supervisor Professor Rolf Larsson without whom I, almost surely, would not have aspired to become a PhD in the first place. I also owe thanks to my assistant thesis supervisor Associate Professor Johan Lyhagen, Professors Emeritus Anders Christoffersson and Anders Ågren, and Professor Adam Taube for their patience with me and my numerous questions, and for many interesting discussions on topics ranging from ill-conditioned matrices to the decomposition of the life and accomplishments of H. Wold. In particular, my thanks goes to my colleague and friend Anders Eriksson; for our endless discussions and arguments on mathematics, statistics and econometrics, for showing a great interest in my research (and strong confidence in its applicability) and for putting sparkle into my LaTeX source. Thanks also to: Camilla (my sofa will never forgive you), Ylva (note to self: wasabi and calvados don't mix), Nicklas (thanks for the squash lessons), Katrin (I hope to see your thesis soon!), Jenny (remember Sesame Street?), Andreas (thanks for all the 'bread'), Tomas ('tres amigos', anyone?) and Joakim (yeah!). Going out on a tangent, I would also like to thank Anna-Maria Lundins Foundation for my conference scholarships, and Jan Wallanders and Tom Hedelius Foundation (W&H) for funding my research these last four years. A special thanks to Karin Påhlson at W&H for being so helpful and cheerful all the time. Thanks to all the administrative staff at the department, especially Ingrid and Eva for taking care of everything W&H related, Gunilla for always being helpful, and L-G for being such a nice guy.

A special thanks to the opponent of my licentiate thesis, Associate Professor Rickard Sandberg, for his many valuable remarks.

The main part of this thesis was written during a one-year stay at the Singapore Management University (SMU). I would herewith like to thank Associate Professor Jun Yu for the invitation and Dean Roberto S. Mariano for the office space, and for an exceptional research environment. Thanks also to my colleagues at SMU for their excellent hospitality. My special thanks goes to Jun Yu for our discussions on financial econometrics and for sharing his experience on how to write research papers.

Thanks to my extended family, my incredible friends, who have supported me in all of my activities. You guys are crazy and I love it! A multitude of thanks to my old friend Henrik Wanntorp for proofreading some of my early work. It has been said that it is easier to square a circle than to get round a mathematician. This is certainly true with him.

My unbounded thanks go to my family, my amazing parents and brothers, who encouraged me so much in life and helped me become the person I am today. You taught me values of determination, patience and hard work, and were always supportive of me in pursuing my dreams. I dedicate this thesis to you. I love you guys! A very special thanks to my grandma for all her support and encouragement along the way.

Finally, I owe thanks to my best friend. You showered me with love, kindness and care. I am lucky to have been blessed with such a wonderful girlfriend. I love you Uma.

Daniel Preve
Singapore, March 26th 2008

Bibliography

[1] F. M. Bandi and J. Russell. Microstructure noise, realized volatility, and optimal sampling. *Working paper, Graduate School of Business, University of Chicago*, 2006.

[2] G.E.P. Box and D.R. Cox. An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, 26(2):211–252, 1964.

[3] W. Chen and R. Deo. Power transformations to induce normality and their applications. *Journal of Royal Statistical Society, Series B*, 66:117–130, 2004.

[4] M. Chernov, A. R. Gallant, E. Ghysels, and G. Tauchen. Alternative models for stock price dynamics. *Journal of Econometrics*, 116:225–257, 2003.

[5] F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):134–145, 1995.

[6] P.D. Feigin and S.I. Resnick. Estimation for autoregressive processes with positive innovations. *Communications in Statistics - Stochastic Models*, 8(3):479–498, 1992.

[7] D. Harvey, S. Leybourne, and P. Newbold. Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13:281–291, 1997.

[8] B. Nielsen and N. Shephard. Likelihood analysis of a first-order autoregressive model with exponential innovations. *Journal of Time Series Analysis*, 24(3):337–344, 2003.

[9] L. Zhang, P. A. Mykland, and Y. Aït-Sahalia. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411, 2005.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 39

Editor: The Dean of the Faculty of Social Sciences

A doctoral dissertation from the Faculty of Social Sciences, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-8638

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2008