

A Low-Dimension Collinearity-Robust Test for Non-linearity

Abstract A new test for non-linearity is developed using weighted combinations of regressor powers based on the eigenvectors of the variance-covariance matrix. The test extends the ingenious test for heteroskedasticity proposed by White (1980), but both circumvents problems of high dimensionality and collinearity, and allows inclusion of cubic functions to ensure power against asymmetry or skewness. Furthermore, the test can handle more non-linear variables than observations. A Monte Carlo analysis compares the performance of the test to the optimal infeasible test and to a variant of White's test. The relative performance of the test is encouraging: the test has the appropriate size and has high power in many situations. Furthermore, collinearity between regressors can increase the power of the test.

Testing for Dynamics in the Conditional Asymmetry: a Residual-based Approach.

Abstract Unconditional and conditional asymmetry is frequently reported when analyzing time series, see Simkowitz and Beedles (1980), Kon (1982), So (1987), French, Schwert, and Stambaugh (1987) and Mittnik and Paoletta (2000), Lambert and Laurent (2001b), Giot and Laurent (2003), Hansen (1994), Harvey and Siddique (1999) to cite a few in the finance literature. Hansen (1994) proposes to estimate directly the conditional asymmetry and kurtosis by extending the ARMA and GARCH ideas to the third and fourth order moments. He relies on a skewed Student distribution that nests the symmetric Student when the asymmetry coefficient equals 0. Dynamics is introduced in the asymmetry by conditioning the asymmetry parameter on past residuals and their squares in a way similar to the conditional variance in an ARCH model. A similar idea has been used by Harvey and Siddique (1999), who use a noncentral-t distribution (scaled to have a unit variance) and condition the third moment on past residuals to the power one and three.

See also Lambert and Laurent (2001b) when the distribution of the innovations is the skewed location-scale family. In this paper, we show how dynamics in asymmetry can be tested (without estimating the full dynamic model by ML as in Hansen (1994), Harvey and Siddique (1999) and Lambert and Laurent (2001b)) when the true conditional distribution belongs to the family of skewed distributions proposed by Fernandez and Steel (1998) and extended by Lambert and Laurent (2001a) and Giot and Laurent (2003).

We assume that a times series model with an appropriate description of the dynamics in the conditional 1st and 2nd order moments has been fitted in a first step on $\psi\tau$ ($\tau = 1; \dots; T$). A skewed fat-tailed distribution with constant asymmetry is assumed for the innovations. It can be generated from a fat-tailed unimodal density using the method proposed by Fernandez and Steel (1998). It shares the same mode and tail properties as the starting density and includes an extra parameter ≈ 2 equal to the ratio of probabilities above and below the mode. This ratio is obviously one for symmetric distributions like the normal or the Student.

Conditionally on the 1st step estimation, one can record at any time τ whether the observed series $\psi\tau$ is below the fitted conditional mode ($\neq\tau$), yielding a time series of indicators that can be related to the conditional skewness parameter $\approx 2\tau$.

Under the assumption of no dynamics in the asymmetry of the distribution, the indicator variable defined as $I\tau = I(\psi\tau < \neq\tau)$, where $I(\cdot)$ is the indicator function, should be uncorrelated with any regressors like its past values, past residuals, day-of-the week dummies, etc. This assumption might be tested by running a simple linear probability model or a logit (probit) model on $I\tau$ including some of the above mentioned information variables as regressors and by examining the statistical significance of the regression parameters.

Alternatively, one can follow Premaratne and Tay (2002) in applying the non-parametric Runs (Wald-Wolfowitz) Test on $I\tau$. As shown by these authors, the assumption of no dynamics in the asymmetry of the distribution implies the randomness of $I\tau$.

The problem of these three tests is that they are built on the residuals of the first step estimation. As a consequence, the usual ordinary least squares (OLS) and ML results do not apply as highlighted by our first Monte Carlo study. The three tests under-rejects the null hypothesis of no dynamics in the asymmetry when dynamics is introduced in the conditional mean (or mode) equation. For instance the size of the three tests is closed to 0 for nominal values of 10, 5 and 1% when an AR(1) term is included in the model (even when the DGP has no AR term).

Using the results in Pierce (1982), we compute the correct asymptotic variance-covariance matrix of the three test statistics by taking into account the uncertainty in the estimation of the parameters in the first-step.

In a comprehensive Monte Carlo study, we show that the size of the test can be much lower than the targeted significance level if the uncertainty in the first-step is ignored. The power can also be very small. A relatively large power and the correct size are obtained when the asymptotic standard errors

are corrected.

Finally, we apply the test on several daily stock index returns and show that some of them do exhibit significant dynamics in the degree of asymmetry (even when there is no significant unconditional skewness).

A Distribution-Free Test for Changes in the Distribution

Abstract This paper considers a new, nonparametric approach to the problem of testing for a structural break. The null hypothesis of no change in the whole distribution of the data is tested using a combination of linear rank statistics. The asymptotic distribution of the test is derived and found to be free from nuisance parameters. The test is found to have nontrivial power versus contiguous alternatives; based on Pitman efficiency, a selection rule is derived for choosing the optimal number of rank statistics to be included in the test.

I(2) Cointegration Analysis in the Presence of Deterministic Shifts

Abstract This paper investigates limit theory for the likelihood analysis of an I(2) cointegrated vector autoregressive (VAR) model subject to deterministic shifts. An I(2) VAR model allowing structural breaks in deterministic terms is introduced, and a likelihood ratio test statistic for integration indices in the I(2) VAR model is then considered. It is shown that the asymptotic distribution of the test statistic can be expressed in the form of a generalised Dicky-Fuller type distribution, so the tabulation of its quantiles is feasible. An empirical analysis of the I(2) VAR model with a deterministic shift is also performed using macroeconomic data in Japan.

Forecast Adjustment and Learning

Abstract Unanticipated breaks can have severe effects on economic forecasts. This paper examines the impact of such shifts on forecasting models, focussing on three themes. First, what information is generated by a forecast model when there is a shift, and how could this be used by agents who want to learn how to improve subsequent forecasts? Secondly, even if agents know that shifts may take place, and be aware of their nature, can they use this knowledge to their advantage? And thirdly, can seemingly arbitrary transformations be used to correct forecasts immediately after the agents detect forecast failure, even though the source of this failure might be unknown? The paper discusses these problems first in a theoretical context, where the success of some atheoretic forecast models is examined, before we move on to an applied example of forecasting in a small model of the UK economy.

Forecasting Good Volatility and Bad Volatility

Abstract The distributions of the returns of many financial assets show significant skewness, but this feature has been only marginally dealt with in literature. Our conjecture is that this asymmetry may be due to a different dynamic behaviour of positive returns with respect to negative ones. Therefore, we propose a GRACH-type model for modelling the possible different dynamics and conditional distributions of positive and negative returns. The stochastic properties of the model are considered and the model is, then, applied to real asset returns and tested against standard GARCH specifications.

Model-based Estimation of High Frequency Jump Diffusions with Microstructure Noise and Stochastic Volatility

Abstract With the advent of high frequency, intraday data, research has been instigated into the intraday and integrated volatility, measured through e.g. realised volatility. Such measures may be contaminated by microstructure effects, or by a jump diffusion, which led to the development of quadratic variation measures. Instead of using such model-agnostic, non-parametric measures, this article focuses on a high-frequency parametric model allowing explicitly for microstructure effects, jump-diffusions, missing observations and stochastic volatility, to get to a parametric measure of volatility to be compared with alternative non-parametric measures. The performance of the model-based estimation on daily variation is compared to the nonparametric approach in both a simulation exercise and an application concerning the Euro/US Dollar exchange rate over the year 2005. The model-based measures are found to perform at least on par with the non-parametric methods, while allowing for further extension taking up macro-economic or other financial variables for modelling jumps in the price process, and thus help in obtaining a better understanding of the inner workings of financial markets.

Forecasting, Structural Breaks, and Non-linearities.

Abstract The talk synthesizes many strands of research in Oxford into modelling and forecasting location shifts - the main source of forecast failures - which also serves as background to my co-authors' presentations. A review of predictability, its properties, and its relation to forecastability highlights the advantages of distinguishing two information sets: 'normal causal factors', and 'forces inducing shifts'. Next, we consider the detection of in-sample location shifts by impulse saturation techniques using *Autometrics*, as impulses are also important when formulating, testing for, and modelling non-linearity.

Many sub-problems need solved en route, including the reduced precision of forecasts after breaks, specifically from changes in collinearity, which provides a case for estimation updating. Finally, as insurance against mis-forecasting breaks, we consider combining different forecasting devices and exploiting their differential robustness to breaks, as well as differencing and intercept corrections to facilitate handling breaks. An application to UK M1 after the Banking Act of 1984 illustrates.

Econometric Modelling When There Are More Variables Than Observations

Abstract Automatic model selection procedures are an increasingly popular tool for the empirical modeller. Among these, general-to-specific methods have been found to outperform forward selection methods. However, increased data collection may make the starting model for general-to-specific modelling unidentified: there are more variables than observations. One reason may be the flexibility in the initial lag structure. Another example, recently considered by Hendry and Santos, is to add a dummy variable for each observation. The objectives of this paper are (1) to make automatic modelling operational under these circumstances, (2) to investigate the performance of the proposed procedures using Monte Carlo methods, (3) to provide a comparison with some simpler forward selection methods, and (4) to consider some new applications of the procedure.

The Impact of Macro News on the Term Structure

Abstract The evaluation of the impact of the news effects is one of the key questions in financial economics and a hot topic in recent studies of macroeconomic analysis. We adopt the automatic algorithm based on general to specific models to handle the large number of economic indicators without limiting the analysis to subjective group of variables. First, we explain the variation in the very short dated US term structure i.e. Fed funds future contracts using macroeconomic surprises. Fed funds interest rates and its expectations are the first link of transmission of Federal Reserve policy to other interest rates. This empirical approach helps to understand the underlying methodological problem and to discuss which model is appropriate. We then show that disagreements are non neutral in the understanding of news surprises. Finally, we measure the impact of two types of endogenous sentiment indicators (accumulated surprises and accumulated disagreement), derived from the analysts surveys conducted for each economic release.

Building Dynamic Marketing Models When There Are More Observations Than Variables

Abstract Typically, marketing departments of modern companies have access to an enormous amount of data related not just to the company itself, but also to the activities of competitors. Traditionally, much of the modeling activity focused on cross-section data. However, with increasingly longer time-series available, dynamic model building becomes of interest. We consider some solutions to the challenges that we faced in the empirical model building. A particular issue that we needed to address was that we had more observations than variables. We investigate the effectiveness of our approach through a parametric bootstrap.

Extracting Business Cycles Using Semi-parametric Time-Varying Spectra with Applications to U.S. Macroeconomic Time Series

Abstract A growing number of empirical studies provides evidence that dynamic properties of macroeconomic time series have been changing over time. Model-based procedures for the measurement of business cycles should therefore allow model parameters to adapt over time. In this paper the time dependencies of parameters are implied by a time dependent sample spectrum. Explicit model specifications for the parameters are therefore not required. Parameter estimation is carried out in the frequency domain by maximising the spectral likelihood function. The time dependent spectrum is specified as a semi-parametric smoothing spline ANOVA function that can be formulated in state space form. Since the resulting spectral likelihood function is time-varying, model parameter estimates become time-varying as well. This new and simple approach to business cycle extraction includes bootstrap procedures for the computation of confidence intervals and real-time procedures for the forecasting of the spectrum and the business cycle. We illustrate the methodology by presenting a complete business cycle analysis for two U.S. macroeconomic time series. The empirical results are

promising and provide significant evidence for the great moderation of the U.S. business cycle. It will be shown that by an effective use of Ox/SsfPack, these complicated time-varying models can be analysed using computationally efficient and fast methods.

Modeling Meteorological Predictors of the Abundance of Deer Mice (*Peromyscus maniculatus*) in the Northwestern United States

Abstract In 1993, the Centers for Disease Control and Prevention (CDC) responded to an outbreak of frequently fatal respiratory distress syndrome in the Four Corners region of the Southwestern United States. The disease, later named hantavirus pulmonary syndrome (HPS), was found to be caused by a previously unknown hantavirus (Sin Nombre virus) carried by the deer mouse (*Peromyscus maniculatus*). HPS is most frequently contracted by inhalation of the aerosolized virus associated with urine, saliva, or nesting materials of infected mice.

To assess the longitudinal epidemiological risk to humans, CDC initiated long-term studies of dynamics of rodent hosts of SNV in the western United States.

A primary objective was to identify the determinants of hantavirus transmission in host populations in order to develop models that could be used to forecasting deer mouse population abundance. Using an abundance measure called “minimum number alive”, time series were collected from trapping arrays in the Western US. This analysis represents our attempt to develop and evaluate an ecological forecasting model of deer mouse population abundance in Montana using available local weather data.

A structural time series model for the natural log of minimum number alive was developed using program STAMP. Optimal time series predictors were selected in-sample with a multi-path general-to-specific search, using PcGets.

Three month forecast horizons were used for each of the four seasons of the year. Conditional forecasts were generated by projecting significant temperature and weather predictor variables into the forecast horizon with ARIMA models and sequentially updating the dependent series of the natural log of MNA with an augmented Kalman filter and a diffuse prior. Forecasts over each season were generated with a sequential inclusion of sets of predictors. The hierarchical inclusion of different sets of variables permitted the testing of the hypothesis that local weather controls deer mice abundance by revealing how much each set of variables contributed to model fit and forecast accuracy.

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention or the Department of Health and Human Services.