Econometrics of Energy Markets

MSc Energy, Finance, Carbon
Paris Dauphine University, Spring 2010

Julien Chevallier*

December 11, 2009

- **Goal:**

  The objective of this course is to provide students with a good command of time-series tools that may be used for the analysis of energy markets. The course covers a wide range of applications in the oil, natural gas, coal, electricity and CO_2 markets. The emphasis is placed on financial econometrics, through the use of linear regression models, vector autoregressive models, cointegration models, and GARCH(p,q) models. Students will be trained with the Eviews softare in computer labs in order to replicate research papers in the field of their choice.

- **Pre-requisites:** Econometrics I and II.

- **Structure:** 18 hours course (equivalent to 2.5 ECTS credits) in computer rooms, divided in 6 lectures of 3 hours. Practical applications under Eviews software. Mondays 01:45PM-05:00PM from January 4 to February 8, 2010. Rooms: A211 Weeks 1-5, A213 Week 6.

- **Assessment:** 50% by group project, 50% by 2-hour written exam.

- **Eligibility:** Students enrolled in the MSc Energy, Finance, Carbon - both in research and professional majors - may register for this course. Students from other MSc need the explicit agreement of the teacher *before* starting the course to be registered. **The course is closed to students outside of Paris Dauphine University**.

*Paris Dauphine University (CGEMP/LEDa), Place du Marechal de Lattre de Tassigny, 75775 Paris Cedex 16, France. *julien.chevallier@dauphine.fr*

1 Lecture #1: Linear Regression Models

Required reading: Weisberg, S. *Applied Linear Regression.* Wiley Series in Probability and Statistics: Chapters #2,3,4,8, 10

1.1 The linear regression model

1.1.1 Presentation

1.1.2 Matrix Notation

1.2 Estimation and Properties of the Estimators

1.2.1 Estimation of regression coefficients

1.2.2 Hypotheses and properties of the estimators

1.2.3 Variance analysis and quality of the adjustment

1.3 Statistical Tests

1.3.1 The role of hypotheses

1.3.2 The construction of tests

1.4 Variance analysis

1.4.1 Variance analysis table and global significance test of the regression

1.4.2 Other tests in variance analysis

1.5 Use of Dummy Variables

1.5.1 Construction and purpose of dummy variables

1.5.2 Examples

1.6 Forecasting

1.6.1 Conditional forecasting

1.6.2 Reliability of forecasting and confidence interval

1.7 Multicolinearity and Variables Selection

1.7.1 Partial correlation

1.7.2 Multicolinearity

1.7.3 Selection of explanatory variables

1.8 Model misspecification

1.8.1 Errors autocorrelation

- Problem identification
- Generalized OLS estimator
- Detecting errors autocorrelation
- Estimation procedures in presence of errors autocorrelation

1.8.2 Residuals stability tests

- Rolling regressions
- CUSUM test

1.8.3 Heteroskedasticity

- Problem identification
- Correcting the heteroskedasticity
- Detecting the heteroskedasticity
- Other tests: the ARCH test

1.9 Problem Set #1

Download the dataset at the following address:

`http://economix.u-paris10.fr/docs/309/data_efc.xls`

1.9.1 Qualitative observation of coal, natural gas, and power spot prices trajectories

Plot the daily spot price trajectories of EEX coal and EEX natural gas prices.

Comment:

- do you observe spikes (upside peaks followed by downside jumps) in coal and/or gas prices?

- are the trajectories formed by random oscillations? do the price series evolve around a random trend?

- does volatility seem higher and spikes more frequent during some specific periods?

Plot the daily spot price trajectory of Powernext EPEX electricity prices.

Comment:

- do you observe a seasonal pattern in the power spot price series? is that different from coal and gas prices?

- can you identify whether downside peaks in power spot price series correspond to bank holidays?

- are the spikes in power spot prices sharper than those observed in other price series? if yes, which one(s)?

- are the trajectories of power spot prices mean-reverting (characterized by random oscillations around a long-term average)?

- does volatility seem higher and spikes more frequent during specific periods?

The seasonality of some commodity spot prices arises from the systematic variations of the demand or supply across hours/days/season and from the impossibility or the difficulty of storing the commodity to absorb these demand variations.

Mean-reversion in commodity prices can be explained by the following mechanisms: when prices are high, producers are urged to use their spare production capacity (inducing short-term reversion in the prices) or to drill new wells (inducing long-term mean reversion in the prices) and consumers tend to decrease their consumption e.g. by switching to substitutes, prompting the prices to go down; when prices are low, producers reduce the quantity put on the market by stopping or storing production and consumers increase their consumption e.g. by switching away from substitutes, which prompts the prices to go up.

1.9.2 Seasonality in Power Prices

Plot the auto-correlations functions of the daily Powernext EPEX spot price..

- Do you observe correlation peaks following a specific pattern?

- Which prices between day d and day $d - n$ seem more correlated than prices with other lags?

Before modelling the random part of the spot prices, it is necessary to filter out the weekly seasonality present in the prices. A simple way of doing this is to:

- compute the average log price L_d for $d =$ days-of-week.

- substract from these average log prices the global average log price

L across all days to obtain an additive seasonality adjustment term $\overline{L_d} = Ld - L$, such that $\sum_{d=1}^{n} \overline{L_d} = 0$

- calculate the seasonality adjustment factors $\pi_d = exp(\overline{L_d})$ such that $\Pi_{d=1}^{n} \pi_d = 1$

- divide the prices of the original series by these adjustment factors to obtain the deseasonalized price series : $\overline{S_t} = \frac{S_t}{\pi_{d_t}}$ where d_t is the week day of date t

Report the adjustment factors for each day of the week.

- when do you observe that prices are the highest (which day-of week)?

Plots the deseasonalized price series, where you have removed the bank holidays.

- Do you still observe any systematic patterns (compared to the previous figure)?

Plot the auto-correlation functions of the two deseasonalized series.

- Do you still observe regular bumps occurring every interval between day d and day $d - n$?

Remark :

- to deseasonalize the hourly spot prices, it is necessary to repeat the above procedure for each of the 24 hourly time series $H1, H2..., H24$

- when a yearly seasonality is present, one can perform an Ordinary Least Square regression of the log prices on sinusoidal functions of period 1 year (in some situations, 6 months and 3 months seasonal patterns must be added) to obtain the deseasonalized data ; however, in many situations, the yearly seasonality is not constant and therefore cannot be well represented by trigonometric series of yearly seasonality : this phenomenon is due to the variability of weather and storage conditions across winters ; in these situations, a wavelet analysis might help decomposing the signal into the sum of a long-term trend, a yearly seasonality with varying shapes (determined essentially by the strength and the starting/ending dates of the winter), and a residual noise which will be modeled as a mean-reverting process.

2 Lecture #2: Basics of Time-Series Econometrics

Required reading: Lutkepohl, H. and Kratzig, M. 2004. *Applied Time Series Econometrics*, Cambridge University Press: Chapter #2.

2.1 Wold's decomposition theorem

2.2 Lag operator

2.3 ARMA Processes

2.3.1 AR and MA processes

2.3.2 ARMA process

2.4 Stationarity

2.4.1 Definition and properties

2.4.2 Simple and partial autocorrelation functions

2.4.3 White-noise tests

2.5 Problems linked to non-stationarity and unit root tests

2.5.1 DS and TS processes

2.5.2 Unit root tests (ADF, PP, KPSS) and sequential test procedures

2.6 Chow's break test

2.6.1 Principle of the test

2.6.2 Sub-samples estimation

2.7 Problem Set #2

2.7.1 Stationarity Tests in Power, Coal and Natural Gas Spot Prices

Statistically, mean-reversion is evidenced by testing the nullity of the coefficient α in the following regression:

$$\Delta x_t = x_{t+1} - x_t = \alpha x_t + \beta + e_t$$

This test is called the Dickey-Fuller test :

- if α is significantly negative, then we say that the process x_t has no unit root, or that it is stationary, inducing a mean-reverting behavior for the prices;

- if α is not significantly different from 0, then we say that the process x_t 'has a unit root', inducing a random walk behavior for the prices.

In practice, Augmented-Dickey-Fuller (ADF) or Phillips-Peron (PP) tests are used rather than Dickey-Fuller. These tests are based on the same principle but corrects for potential serial autocorrelation and time trend in Δx_t through a more complicated regression :

$$\Delta x_t = \sum_{i=1}^{L} \beta_i \Delta x_{t-i} + \alpha x_t + \beta_1 t + \beta_2 + e_t$$

The ADF test tests the null hypothesis H_0 that č $\alpha = 0$ (the alternative hypothesis H_1 being that $\alpha < 0$) by computing the Ordinary Least Squares (OLS) estimate of α in the previous equation and its t-statistics $\hat{t}$; then, the statistics of the test is the t-statistics $\hat{t}_\alpha$ of coefficient α, which follows under H_0 a known law (studied by Fuller and here denoted Ful). The test computes the p-value p, which is the probability of $Ful \leq \hat{t}$ under H_0. If

$p < 0.05$, H_0 can be safely rejected and H_1 accepted: we conclude that the series 'x_t has no unit root'.

Extensions of these stationarity tests were also developed by Phillips and Perron (PP, 1988) and Kwiatkowski, Phillips, Schmidt, and Shin (KPSS, 1992).

Report the ADF, PP, and KPSS tests for logarithmic Powernext electricity, Powernext gas, EEX gas and EEX coal spot prices.

- What do you conclude? For which time-series can H_0 be rejected? For which time-series can you not reject H_0?

- Detail the methodology to transform non-stationary variables to stationary, and the possible bias induced if the process is TS or DS.

2.7.2 Calibration of Mean-Reversion Models

The simplest mean-reversion model is the so-called AR(1) model:

$$x_{t+1} = \rho x_t + \beta + e_{t+1}$$

where the (e_t) are i.i.d with law $N(0, \sigma^2)$.

The calibration proceeds in two steps :

1. the coefficients ρ and β are determined by OLS regression of x_{t+1} on x_t;
2. σ is computed as the standard deviation of the residuals of the regression

Calibrate this model to the electricity, coal and gas log spot prices.

The mean-reversion characteristic time is computed as $\tau = \frac{1}{1-\rho}$. It equals 15 days for electricity, 30 days for coal and gas. We are faced now with four types of problem using the mean-reversion model :

1. the residuals of the AR(1) model are correlated (**test the hypothesis H_0 that the residuals are uncorrelated using the Box-Pierce test**): this could be dealt with by switching to the ADF model;

2. the residuals do not follow a normal distribution (**test the hypothesis H0 that the residuals follow a Gaussian law; this test is called the Jarque-Bera test**). This difficulty could be overcome by:

   - replacing the normal distribution by a density presenting fat tails (e.g. the Normal Inverse Gaussian or Generalized Error Distribution), but in this case, nothing would guarantee that large positive jumps will first occur and then be followed by large negative jumps;

   - by introducing the possibility of random positive jumps occurring at random times (the negative 'jumps' being generated by the mean-reversion parameter $\alpha = \rho - 1$:

   $$x_{t+1} - x_t = \alpha x_t + \beta + e_{t+1} + q_{t+1} J_{t+1}$$

   where q_{t+1} is a random variable taking values 1 (jump) or 0 (no jump) with probabilities p and $1-p$ and J_{t+1} is a positive random jump; but in this case, the mean-reversion speed α will be too high to account for mean-reversion in normal periods as it has to capture the mean-reversion in both normal periods and spike regimes.

3. the residuals present heteroscedasticity: their variance is not constant in time (**test the hypothesis H_0 that the squared residuals are uncorrelated using the Box-Pierce test**); this could be overcome by introducing a GARCH (Generalized Auto Regressive Conditional Heteroscedasticity) model (see Lecture 5).

To visualize these difficulties, **plot the graphs testing the relevance of AR(1) model on electricity, coal and gas log prices.**

A second test of relevance for a model (M) is to **compare the observed trajectories with the ones generated by the model (M)**.

- What do you observe? Which simulated path under the AR(1) model with normal noise does not look like the observed one (absence of peaks, too large volatility compared to normal periods, etc.)?

- Which simulated path looks more like the observed path, apart from the absence of peaks and the non-stochastic and non-seasonal volatility?

Assess the realism of the paths generated by the AR(1) model (with NIG noise) and the random positive jumps model for coal natural gas.

Plot the observed and simulated paths under the random positive jumps model with reasonable choice of parameters.

- What do you observe? Which model produces much more realistic paths than the simple auto-regressive the AR(1) model?

Note that to produce fast mean-reverting jumps, the parameter α has to be highly negative, provoking excessive mean-reversion speed in normal periods.

3 Lecture #3: Vector AutoRegressive Models

Required reading: Watson, M.W. 1994. Vector autoregressions and cointegration. *Handbook of Econometrics* 4, 2843-2915.

3.1 VAR Model Representation

3.1.1 Example

3.1.2 General representation

3.1.3 ARMAX representation

3.2 Parameters estimation

3.2.1 Estimation method

3.2.2 Lag determination

3.2.3 Forecasting

3.3 VAR Dynamics

3.3.1 VMA representation

3.3.2 Shocks analysis: Impulse Response Functions

3.3.3 Variance decomposition

3.4 Causality

3.4.1 Causality in the Granger sense

3.4.2 Causality in the Sims sense

3.5 Problem Set #3

Dependence models for stationary commodity prices

The traditional model for correlated stationary prices is the Vector-Auto-Regressive model (VAR), which is the multi-variate extension of the Auto-Regressive model.

Let $Z_t = \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix}$ be the vector process formed of the two (properly deseasonalized) log prices. Then, the VAR(p) model reads :

$$Z_t = C + \Gamma_1 Z_{t-1} + \ldots + \Gamma_p Z_{t-p} + \epsilon_t$$

where $C = \begin{pmatrix} C^1 \\ C^2 \end{pmatrix}$ is a constant vector, $\Gamma_1, \ldots, \Gamma_p$ are 2x2 matrices and the vector process $\epsilon_t = \begin{pmatrix} \epsilon_t^1 \\ \epsilon_t^2 \end{pmatrix}$ is formed of independent random variables following a centered bi-variate normal distribution $N(0, \sum)$.

The calibration of the model VAR(p) proceeds in three classic steps:

1. the optimal order p is selected using an information criterion, which is an indicator of the relevance of a model, giving a positive weight to the likelihood of the model and a negative weight to the number of model parameters ; e.g., the Schwartz Information Criterion (SIC) is equal to $2LL - ln(T)n$, where LL is the log-likelihood, T is the number of observations and n is the number of parameters of the model.

2. once p is known, $C, \Gamma_1, \ldots, \Gamma_p$ are determined by OLS.

3. lastly, the standard deviation and correlation of the residuals give matrix $\sum$

Perform the calibration of the VAR model on the pairs Powernext electricity/gas and EEX gas/coal.

Report the following results:

1. the lag order optimizing the information criteria;

2. the results of the two OLS regressions on lagged prices;

3. the correlation between the electricty and natural gas residuals.

Assess whether the causality runs positively (negatively) from Powernext electricity (natural gas) to natural gas (electricity), and explain why with respect to the regression coefficients of electricity (natural gas) on the lagged natural gas (electricity) prices (sign & significance).

Repeat the analysis with EEX gas/coal prices.

More generally, we will say that a process P_t^1 Granger causes P_t^2 at the order p if, in the linear regression of P_t^2 on lagged prices $P_{t-1}^1, \ldots, P_{t-p}^1, P_{t-1}^2, \ldots, P_{t-p}^2$, at least one of the regression coefficients of P_t^1 on the lagged prices $P_{t-1}^2, \ldots, P_{t-p}^2$ is significantly different from 0. The intuition behind Granger causality is that the information on past prices $P_{t-1}^2, \ldots, P_{t-p}^2$ is relevant to forecast P_t^1 at future time t.

Granger causality is examined using the Granger causality test testing the null hypothesis H_0 that all regression coefficients of P_t^1 on the lagged prices $P_{t-1}^2, \ldots, P_{t-p}^2$ are null. A p-value lower than 0.05 means that H_0 can be rejected (and causality accepted) with 95% confidence level.

Perform the Granger causality tests on the pairs Powernext electricity/gas and EEX gas/coal.

How well does the VAR(p) model perform to account for the spikes observed in the trajectories of previous problem sets?

A model capturing correlated jumps between commodity spot prices is proposed by Benth and Kettler (2006) : the idea consists in describing the gas and electricity log spot prices by a VAR(1) process with non-Gaussian correlated residuals modeled by a copula representation.

The idea behind copulas is to disentangle the problems of fitting the margin distributions and the dependence structure. Typically, the first task is performed first (using e.g. Normal Inverse Gaussian distribution for each residual separately), followed by the maximum-likelihood fit of a chosen class of copula.

Perform IRF analysis based on the most appropriate choice of decomposition for the pairs Powernext electricity/gas and EEX gas/coal.

Perform variance decomposition analysis for the pairs Powernext electricity/gas and EEX gas/coal.

4 Lecture #4: Cointegration and Error Correction Models

Required reading: Watson, M.W. 1994. Vector autoregressions and cointegration. *Handbook of Econometrics* 4, 2843-2915.

4.1 Examples

4.2 The concept of cointegration

4.2.1 Properties of the order of integration of a time-series

4.2.2 Conditions for cointegration

4.2.3 The Error Correction Model

4.3 Cointegration between 2 variables

4.3.1 Cointegration test between 2 variables: : the Engle-Granger procedure

4.3.2 Estimation of the Error Correction Model

4.4 Cointegration between k variables

4.4.1 Cointegration test between k variables: the Johansen procedure

4.4.2 Estimation of the Error Correction Model

4.4.3 Dynamics and Vector Error Correction Models

4.4.4 Cointegration Relation Tests

4.4.5 Summary of the test procedure

4.5 Problem Set #4

4.5.1 Dependence models for non-stationary commodity prices

For non-stationary processes, it is impossible to use a VAR(p) model, which intrinsically applies to mean-reverting time series. Therefore, we use the concept of cointegration :we look for a stationary linear combination of the two series, which will represent the long-run equilibrium and we study the error-correction mechanisms insuring the reversion to the long-run equilibrium. For example, when we study the correlation between crude oil and gasoline, it is reasonable to test whether the crack-spread is mean-reverting and to expect a correction of an abnormally high/low crack spread via crude oil or gasoline prices.

Let us apply this idea to the pair of EEX gas/Powernext gas spot prices, for which you have studied stationarity in Problem Set #2.

Report the results of the OLS regression of EEX gas spot log prices on Powernext gas spot log prices.

Apply the 3 standard stationarity tests (ADF, PP, KPSS) to the residuals of the regression.

- which obtain p-value do you obtain for the Augmented-Dickey-Fuller test and for the Phillips-Perron test?

- can you safely reject the hypothesis of existence of a unit root, hence conclude to the stationarity of the residuals?

The last step of the cointegration model consists in describing the dynamics of the two series in terms of the residuals of the long-term relation:

$$\begin{pmatrix} \Delta X_t^e \\ \Delta X_t^{e'} \end{pmatrix} = \begin{pmatrix} \mu_e \\ \mu_{e'} \end{pmatrix} + \sum_{k=1}^{p} \Gamma_k \begin{pmatrix} \Delta X_{t-k}^e \\ \Delta X_{t-k}^{e'} \end{pmatrix} + \begin{pmatrix} \Pi_e \\ \Pi_{e'} \end{pmatrix} R_t + \begin{pmatrix} \epsilon_t^e \\ \epsilon_t^{e'} \end{pmatrix}$$

where

- e stands for EEX gas, and e' stands for Powernext gas;

- X_t^e is the log spot price of energy e at time t;
- the 2x1 vector process $\Delta Z_t = \left(\Delta X_t^e = X_{t+1}^e - X_t^e, \Delta X_t^{e'} = X_{t+1}^{e'} - X_t^{e'}\right)'$ is the vector of EEX gas and Powernext gas price returns;
- $\mu = (\mu_{X,e}, \mu_{X,e'})$ is the 1x2 vector composed of the constant part of the drifts;
- Γ_k are 2x2 matrices expressing dependence on lagged returns;
- $(R_t = X_t^e - \beta X_t^{e'})$ is the process composed of the deviations to the long-term relation between the EEX gas and Powernext gas log spot prices;
- Π is a 2x1 vector matrix expressing sensitivity to the deviations to the long-term relation between the EEX gas and Powernext gas prices;
- the residual shocks $(\epsilon_t^e, \epsilon_t^{e'})$ are assumed to be i.i.d with a centered bi-variate normal distribution $N(0, \sum)$.

Calibrate the model by following the following steps:

1. check, using unit root tests, that the log prices of EEX gas and Powernext gas are non stationary and integrated of order one: this amounts to checking that they are difference stationary, i.e. ΔX_t^e and $\Delta X_t^{e'}$ are stationary;
2. check that they are cointegrated, i.e. β exists such that $R_t = X_t^e - \beta X_t^{e'}$ is stationary. This can be done by performing an OLS regression of X_t^e on $X_t^{e'}$ or more rigorously by using the Johansen cointegration test;
3. then, the optimal lag p is selected using an information criterion;

4. compute Γ and Π by OLS regression of the returns on lagged returns and past deviation (R_t).

5. the standard deviation and correlation of the residuals $(\epsilon_t^e, \epsilon_t^{e'})$ give matrix $\sum$

- concerning step #4: do you observe that either EEX gas or Powernext gas returns correct the deviations to the long-term equilibrium?

- can you conclude that EEX gas or Powernext gas is the leader in the long-term price discovery?

- can you identify a positive causality runs from EEX gas returns to Powernext gas returns (or conversely)?

- do you observe that EEX gas price evolution is completely independent of Powernext gas prices?

4.5.2 VECM Analysis

We want to introduce an error-correction mechanism on the levels and on the slopes between the energies e and e'.

Let X_t be a vector of N variables, all $I(1)$:

$$X_t = \Phi_1 X_{t-1} + \ldots + \Phi_p X_{t-p} + \epsilon_t$$

with $\epsilon_t \sim WGN(0, \Omega)$, WGN denotes the White Gaussian Noise, Ω denotes the variance covariance matrix, and Φ_i $(i = 1, \ldots, p)$ are parameter matrices of size (NxN).

Under the null H_0, there exists r cointegration relationships between N variables, *i.e.* X_t is cointegrated with rank r.

The error correction model writes:

$$\Delta X_t = \Pi_1 \Delta X_{t-1} + \ldots + \Pi_{p-1} \Delta X_{t-p+1} + \Pi_p X_{t-p} + \epsilon_t$$

where the matrices Π_i $(i = 1, \ldots, p)$ are of size (NxN).

All variables are $I(0)$, except X_{t-p} which is $I(1)$.

For all variables to be $I(0)$, $\Pi_p X_{t-p}$ needs to be $I(0)$ as well.

Let $\Pi_p = -\beta\alpha'$, where α' is an (r, N) matrix which contains r cointegration vectors, and β is an (N, r) matrix which contains the weights associated with each vector.

If there exists r cointegration relationships, then $Rk(\Pi_p) = r$. Johansen's cointegration tests are based on this condition.

$$\Delta X_t = \Pi_1 \Delta X_{t-1} + \ldots + \Pi_{p-1} \Delta X_{t-p+1} - \beta\alpha' X_{t-p} + \epsilon_t$$

Estimate the VECM through maximum likelihood methods.

Compute the trace test statistics and maximum eigenvalue test statistics associated with Johansen's methodology.

Plot two examples of simulated price paths under the VECM model.

- What do you observe from these two graphs?

- Does the VECM model generate drastically different long-term price behaviors?

When evaluating the risk of multi-commodity spot exposures, it is important to filter out the economically non-plausible paths. It is therefore crucial to complement a statistical model with some economic insights giving e.g. some deterministic reasonable bounds for gas and coal prices in the long-term. Then, a constrained VECM model can be simulated with reflecting upside and downside (possibly time-dependent but deterministic) barriers representing the economic upper and lower bounds for commodity prices in the long-term.

5 Lecture #5: GARCH(p,q) Models

Required reading: Bollerslev, B., Engle, R.F., Nelson, D.B. 1994. ARCH models. *Handbook of Econometrics* 4, 2959-3038.

5.1 General presentation and problem identification

5.2 ARCH models

5.2.1 Model specifications

5.2.2 Properties of the ARCH(1) model

5.3 ARCH Model tests

5.4 Estimation method and forecasting

5.5 GARCH Processes

5.5.1 Model specification

5.5.2 Test and estimation method of GARCH models

5.6 Other GARCH models

5.6.1 ARCH-in-mean

5.6.2 EGARCH

5.6.3 TGARCH

5.6.4 IGARCH

5.7 Problem Set #5

Download CO_2 futures prices of maturity December 2010 at the following address:

`http://www.ecx.eu/Market-Data/`

5.7.1 Calibrating GARCH models for carbon prices

Pre-estimation

Examine the descriptive statistics for CO_2 price series: how well do they seem to fit GARCH(p, q) models specifications?

Examine the plots of autocorrelation, partial autocorrelation, and autocorrelation of squared logreturns for the CO_2 price series.

- What do you observe with respect to correlation in the raw returns **(compute the p-value of the Ljung-Box test at 5% level)**, serial correlation in the squared returns **(compute the p-value of the Ljung-Box test at 5% level)**, evidence of heteroskedasticity **(compute the p-value of the ARCH test 5% level)**?

GARCH(p, q) estimation

1. **Using the Box-Jenkins methodology, configure the ARMA(p, q) processes that provide the best fit to the CO_2 time-series.**

2. **Estimate the corresponding GARCH(p, q) model for the CO_2 price series.**

$$Y_t = \theta X_t' + \epsilon_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

with σ_t^2 the conditional variance, which is function of a constant term ω, the ARCH term ϵ_{t-i}^2, and the GARCH term σ_{t-j}^2.

Comment on:

- the statistical significance of the parameters obtained in the mean and variance equations;
- the sensitivity of the results obtained to the estimation methodology (usually QML and BHHH);
- the value of the Ljung-Box test;
- the value of the ARCH test;
- the plot of autocorrelation of squared logreturns for the CO_2 price series;
- the innovations, conditional standard deviations, and standardized innovations of the GARCH (p, q) model.

You should observe:

- No correlation.
- Little volatility clustering. Modeling similar to Benz and Truck (2009).
- Including a dummy variable for the structural break from April 25 to June 23, 2006 is not significant.

Other GARCH(p, q) models

Consider other models than the standard GARCH(p, q) model with a Gaussian conditional probability distribution.

Let us recall the specification for the conditional variance proposed by Nelson (1991):

$$log(\sigma_t^2) = \omega + \sum_{i=1}^{p} \alpha_i \left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| + \sum_{j=1}^{q} \beta_j log(\sigma_{t-j}^2) + \sum_{k=1}^{r} \gamma_k \frac{\epsilon_{t-k}}{\sigma_{t-k}}$$

where γ tests for the presence of the leverage effect[1].

The PARCH specification allows to model the standard deviation rather than the variance, and may be written as follows (Ding et al. (1993)):

$$\sigma_t^\delta = \omega + \sum_{i=1}^{p} \alpha_i (|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i})^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta$$

with $\delta > 0$, $|\gamma_i| \leq 1 \; \forall i = 1, \ldots, \tau$, $\gamma_i = 0 \; \forall i > \tau$, $\tau < p$. δ is the power parameter of the standard deviation, and γ parameters to capture asymmetry up to order τ.

Finally, the asymmetric TGARCH model by Zakoian (1994) may be written as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{k=1}^{r} \gamma_k \epsilon_{t-k}^2 \Gamma_{t-k}$$

where $\Gamma_t = 1$ if $\epsilon_t < 0$, and 0 otherwise. $\epsilon_{t-i} > 0$ and $\epsilon_{t-i} < 0$ denote, respectively, good and bad news.

Do you observe any improvement with GJR-GARCH, EGARCH, PARCH or alternative specification structures in the variance equation based on likelihood ratio tests, AIC and BIC?

5.7.2 Multivariate analysis

Include other energy variables in your analysis (e.g., coal, gas, ...).

Compute the matrix of cross-correlations between energy variables: can you detect some problematic multicollinearities?

[1] Recall that the leverage effect implies a higher level of volatility associated to decreasing prices in the financial economics literature.

Report the GARCH(p, q) estimates for the CO_2 with coal and/or natural gas regressors.

Which variables are significant? Drop non-significant variable to obtain your reduced form model.

How can you interpret the sign of the coefficients obtained?

Provide the plot of autocorrelation of squared logreturns, innovations, conditional standard deviations and standardized innovations, along with usual diagnostic tests.

6 Lecture #6: Group Project

6.1 Guidelines

The aim of the group project (**maximum 3 students by group**) is to replicate a research paper in the field of energy markets. It can be applied to oil, natural gas, coal, electricity and CO_2 markets.

Two priority fields of applications are given to students in the Syllabus:

1. on the *oil* market: replicate key papers on the link between the macroeconomic environment and the oil market, or more recent papers on the role of speculation on the oil market;

2. on the CO_2 market: identify and replicate key research questions in recent working papers.

Besides these two themes, students are free to choose another area of application in the field of energy markets.

Each project needs to be explicitly agreed upon with the teacher.

During this course, students need to identify key research questions - in adequation with technical feasibility and under tight time constraints - and submit their research projects (by email to the teacher for the data used, printed version to Dominique Charbit) no later than **March 8, 2010 by 12:00AM**.

The project will count towards 50% of the mark for this course.

6.2 Philosophy of the project: the end of *Kindergarten*

Especially at the MSc level, this research project should be carried out by students in complete **AUTONOMY**. The teacher will provide supervision to students during Lecture 6, but the students need to be aware of the following steps that are required to complete successfully the project:

1. identify ONE precise research question;
2. mobilize the adequate econometrics tools;
3. review the relevant literature for this research question;
4. gather the dataset;
5. clean the dataset;
6. make the necessary computational steps in order to derive the variables of interest in your project;
7. carry out your econometric analysis;
8. interpret the results obtained;
9. if time allows, develop sensitivity tests for the results obtained;
10. generalize the results obtained in the context of a broader research topic;
11. identify areas for future research.

The amount of time estimated to complete the project lies **between 30 and 60 hours of personal work.** (more especially for *research* students).

6.3 Time-Management

- Steps 1 to 7 typically represent 70% of the time that you will spend on your group project.
- Thus, you need to start as early as possible to gather your dataset and construct your variables once your project has been clearly defined and agreed with the teacher.

6.4 Discussion

- You can do very well in this group project by tackling a simple research question, and by providing a detailed, insightful, intelligent discussion of the results obtained.
- Conversely, you might want to tackle a quite technical research question and fail to achieve any meaningful result.
- **Thus, the emphasis for the evaluation of your group project is clearly placed on testing your ability to complete a research project in due time, with learning effects and under pressing time constraints** (which is in short what happens in real life).

I do not expect you to replicate state-of-the-art econometric techniques (with the notable of research students interested into pursuing Phd studies).

7 Assessment

50% by group project, 50% by 2-hour written exam.

7.1 Group project

Group projects are due on **March 8, 2010**, to Dominique Charbit's office, by 12:00AM. **No late submission will be accepted, and the grade of zero will be automatically placed on late submitted projects.**

Send the Eviews worksfile (or data/code if you use another software for replicability) by email to:

`julien.chevallier@dauphine.fr`

7.2 Written exam

The 2-hour written exam is scheduled on March 8, 2010 from 8:30AM to 10:30AM, room #B315. The emphasis is placed on the understanding of the econometric tools studied in Lectures 1 to 5.

<u>No documents or calculator will be allowed</u>.

Syllabus

Textbooks

- Hamilton, J.D. 1996. *Time Series Analysis*, Princeton University Press.
- Lutkepohl, H. and Kratzig, M. 2004. *Applied Time Series Econometrics*, Cambridge University Press.
- Vogelvang, B. 2005. *Econometrics: Theory & Applications With Eviews*, Pearson Education Limited.
- Weisberg, S. *Applied Linear Regression*. Wiley Series in Probability and Statistics.

Oil Market

Bernanke, B.S., Gertler, M., Watson, M. 1997. Systematic Monetary Policy and the Effects of Oil Price Shocks, *Brookings Papers on Economic Activity* 1, 91-157.

BOE, 2008. What can be said about the rise and fall in oil prices? *Bank of England Research and Analysis Report*, prepared by Saporta, V., Trott, M., et Tudela, M. 215-225.

Buyuksahin, B., Haigh, M.S., Harris, J.H., Overdahl, J.A., and Robe, M.A. 2008. Fundamentals, Trader Activity and Derivative Pricing. *SSRN Working Paper* #966692.

Cifarelli, G., Paladino, G. 2010. Oil price dynamics and speculation: a multivariate financial approach. *Energy Economics*, forthcoming.

Hamilton, J.D. 1996. This is what happened to the oil-price macroeconomy relationship, *Journal of Monetary Economics* 38(2), 215-220.

Kaufmann, R.K., Ullman, B. 2009. Oil prices, speculation, and fundamentals: Interpreting causal relations among spot and futures prices. *Energy Economics* 31, 550-558.

Medlock III, K.B., Jaffe, A.M. 2009. Who is in the oil futures market and how has it changed? *Working Paper*, Rice University, USA.

Parsons, J.E. 2009. Black Gold & Fool's Gold: Speculation in the Oil Futures Market. *MIT CEEPR Working Paper* #09-013.

Tang, K., Xiong, W. 2009. Index Investing and the Financialization of Commodities. *Working Paper*, Princeton University, USA.

CO_2 Market

Alberola, E. and Chevallier, J. (2009). European Carbon Prices and Banking Restrictions: Evidence from Phase I (2005-2007). *The Energy Journal* 30(3), 51-80.

Alberola, E., Chevallier, J. and Chèze, B. (2008) Price drivers and structural breaks in European carbon prices 2005-2007. *Energy Policy*, 36 (2), 787-797.

Alberola, E., Chevallier, J. and Chèze, B. (2009a). The EU Emissions

Trading Scheme: the Effects of Industrial Production and CO2 Emissions on European Carbon Prices. *International Economics* 116, 93-126.

Alberola, E., Chevallier, J. and Chèze, B. (2009b). Emissions Compliances and Carbon Prices under the EU ETS: A Country Specific Analysis of Industrial Sectors. *Journal of Policy Modeling* 31(3), 446-462.

Benz, E., Truck, S., 2009. Modeling the price dynamics of CO_2 emission allowances. *Energy Economics* 31 (1), 4-15.

Chevallier, J. 2009. Carbon Futures and Macroeconomic Risk Factors : A view from the EU ETS, *Energy Economics* 31(4), 614-625.

Chevallier, J., Ielpo, F. and Mercier, L. 2009. Risk Aversion and Institutional Information Disclosure on the European Carbon Market: a Case-Study of the 2006 Compliance Event. *Energy Policy* 37(1), 15-28.

Daskalakis, G., Psychoyios, D., Markellos R. N., 2009. Modeling CO_2 emission allowances prices and derivatives: Evidence from the European trading scheme. *Journal of Banking and Finance* 33, 1230-1241.

Frunza, M.C. and Guegan, D. 2009. An economic view of carbon allowances market. *Working Paper* #2009.38, Centre d'Economie de la Sorbonne.

Gronwald, M. and Ketterer, J. 2009. Evaluating Emission Trading as a Policy Tool - Evidence from Conditional Jump Models, *CESifo Working*

Paper #2682.

Keppler, J.H. and Mansanet-Bataller, M. Causalities between CO_2, Electricity and other Energy Variables during Phase I and II of the EU ETS. *Mission Climat Working Paper* #2.

Oberndorfer, U. 2009. EU Emission Allowances and the stock market: Evidence from the electricity industry. *Ecological Economics* 68, 1116-1126.

Paolella, M. and Taschini, L. 2008. An Econometric Analysis of Emission Allowances Prices. *Journal of Banking and Finance* 32, 2022-2032.

Violante, F. and Sanin, M.E. 2009. Understanding volatility dynamics in the EU-ETS market: lessons from the future. CORE Discussion Paper #2009-24.

Other references

Benth F. and Kettler P. 2006. Dynamic copula models for the spark spread. *Working Paper*, Department of Mathematics, University of Oslo.

Ding, Z., Granger, C.W.J., Engle, R.F. 1993. A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance* 1, 83–106.

Geman H. 2005. Energy Commodity Prices: Is Mean-Reversion Dead?, *Journal of Alternative Investments* 8(2), 31-45.

Geman H. and Ohana S. 2005. Forward Curves, Scarcity, and Price Volatility in Oil and Natural Gas Markets. *Energy Economics* 31(4), 576-585.

Geman H. and Ohana S. 2009. Time-consistency in managing a commodity portfolio: A dynamic risk measure approach. *Journal of Banking and Finance* 32(10), 1991-2005.

Nelson, D.B. 1991. Conditional Heteroskedasticity in Asset Retuns: A New Approach. *Econometrica* 59(2), 347–370.

Ohana, S. 2010. Modeling global and local dependence in a pair of commodity forward curves with an application to the US natural gas and heating oil markets. *Energy Economics*, forthcoming.

Ohana, S. 2008. Quantitative Methods in Commodity Finance. *Unpublished Manuscript*, Birbeck College London.

Zakoian, J.M., 1994. Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control*, 18 (5), 931–944.