Volatility Modelling of CO_2 Spot Prices

Master Thesis Submitted to

Prof. Dr. Brenda López Cabrera

Prof. Dr. Wolfgang K. Härdle

Ladislaus von Bortkiewicz Chair of Statistics

C.A.S.E.- Center for Applied Statistics and Economics

Humboldt-Universität zu Berlin

by

Thijs Benschop

(541993)

in partial fulfillment of the requirements

for the degree of

Master of Science in Statistics

Berlin, March 4th, 2013

Abstract

In this paper we analyse the short-term spot price of European Union Allowances (EUAs), which is of particular importance in the transition of energy markets and for the development of new risk management strategies. We use daily spot market data from the second trading period of the EU ETS. Emphasis is given to short-term forecasting of prices and volatility. Due to the characteristics of the price process, such as volatility modelling, breaks in the volatility process and heavy-tailed distributions, we investigate the use of Markov switching GARCH (MS-GARCH) models. We find that these models distinguish well between states, and that the volatility processes in the states are clearly different. Our findings support the use of MS-GARCH models for risk management, especially because their forecasting ability is better than other Markov switching or simple GARCH models.

Keywords: CO_2 Emission Allowances, CO_2 Emission Trading, Spot Price Modelling, Markov Switching GARCH Models, Volatility Forecasting

List of abbreviations

ACF	autocorrelation function
AIC	Akaike information criterion
ADF	Augmented Dickey Fuller
AR	autoregressive
ARCH	autoregressive conditional heteroskedastic
CO_2	carbon dioxide
DE	Differential Evolution
EUA	European Union Allowance
EU ETS	European Union Emissions Trading System
GARCH	generalized autoregressive conditional heteroskedasticity
GHG	greenhouse gases
KPSS	Kwiatkowski–Phillips–Schmidt–Shin
MAE	mean absolute error
MS	Markov switching
MSE	mean squared error
OTC	over-the-counter
PACF	partial autocorrelation function

Contents

List of figures

List of tables

1 Introduction

It is widely agreed among scientists, politicians and the broader public that the emission of greenhouse gases (GHGs) by human activity has led to an increase in the level of GHGs in the atmosphere, to global warming and climate change. These phenomena have serious impact on the environment, human beings and the economy. In response to these developments many industrialised countries agreed in the United Nations Framework Convention on Climate Change to stabilise the emission of GHGs and adopted the Kyoto protocol, thus accepting its binding obligations to reduce GHG emissions. The Kyoto protocol entered into force in 2005. The member states of the European Union decided to fulfil their commitments jointly and implement a trading system for emission allowances, i.e. permissions to emit one ton of CO_2 in the atmosphere, as a main mechanism to reduce emissions. The European Union Emissions Trading System (EU ETS) entered into force in 2005.

Since the introduction of the EU ETS a new market for European Union Allowances (EUAs) and their derivatives developed and with it carbon finance, a new field of applied econometrics, which investigates the behaviour of prices. The price dynamics and its determinants are of great importance for participating industries, and for sound risk management and hedging strategies of financial intermediaries as well as for policy makers who use them to evaluate the performance of the EU ETS. Furthermore, the market for EUAs is constantly growing, which makes it important for market participants to have a valid pricing model.

Having particular characteristics the EUAs should be regarded as a new class of assets (Benz & Trück, 2006). Under the cap-and-trade scheme of the EU ETS, the total number of allowances is fixed every year and thus the prices are induced by current demand. This is unlike for instance a company stock, where the value is based on profit expectations of the company. The demand is governed by shocks, such as temperature changes, the level of economic activity and energy prices as well as news releases concerning regulatory policy. All these events can alter the production of CO_2 and hence the short-term demand for EUAs.

Since the start of EUA trading a number of studies have focussed on the price determinants of EUA spot prices (e.g. Mansanet-Bataller et al., 2007; Alberola et al., 2007, 2008a,b; Chevallier, 2009; Hintermann, 2010). Focussing on long term relationships, these studies found a relationship between EUA spot prices and energy prices, extreme weather events and economic activity. However, as the relationship changes over time, it depends on the sample and the time period under consideration. Because we are interested in short-term price modelling and forecasting, we do not incorporate externalities in our model. Furthermore, we seek to endogenise the break points by looking for models that fit longer time series. Only a few studies investigate the stochastic behaviour of short-term spot prices and provide an econometric analysis, such as Paolella & Taschini

(2008), Seifert et al. (2008), Daskalakis et al. (2009) and Benz & Trück (2009). The latter investigate the performance of GARCH and Markov regime switching models and find that both approaches give satisfying results.

The EU ETS has had three trading periods so far, the first of which ran from 2005 until 2007. Most available empirical research uses data from this first, pilot, period. The price signals in this period were distorted due to an oversupply of EUAs, which is why we apply data from the second trading period in this study.

Our aim is to analyse the short term spot price behaviour of EUAs traded under the EU ETS by modelling the price movements and the underlying stochastic processes. Some authors (e.g. Paolella & Taschini, 2008) suggest to investigate the performance of regime switching models with GARCH processes in the regimes. The Markov regime switching GARCH (MS-GARCH) model as introduced by Hamilton & Susmel (1994) combines the strength of a regime switching model, which can capture breaks and non-linearities in the underlying stochastic process, with the possibility to model conditional volatility and volatility clustering. The observed breaks and volatility clustering in the EUA spot market suggest that it is worthwhile investigating the application of MS-GARCH models.

In this paper we study the application of the MS-GARCH model as specified by Klaassen (2002) to the log returns of the EUA spot market prices from the second trading period of the EU ETS from 2008 until 2012. To the best of our knowledge this is the first paper to fit MS-GARCH models to EUA log returns. This approach has never been applied because the available time series were too short to estimate such models. In order to compare the performance of the models, we also use simple GARCH and Markov switching models with other state specifications as already investigated by Benz & Trück (2009), who used over-the-counter market data from the first trading period. In this way we can investigate whether their findings also hold when using spot market data from the second trading period. We focus on the forecasting performance of the models as this is the most important aspect for risk management and value-at-risk calculations. We find that MS-GARCH models provide a better in-sample fit and density forecasts. Furthermore, they solve the problem of volatility persistence observed when using simple GARCH models.

The remainder of this paper is organized as follows. Section 2 gives a brief overview of the EU ETS, the EU carbon market and the characteristics of EUAs and their price determinants. Section 3 presents recent literature on the modelling of EUA prices. In section 4 we describe the models used, especially the specification of the MS-GARCH model and estimation procedures. Section 5 provides an empirical analysis of EUA spot market prices and gives the results of the evaluation of the forecasting ability of the models under consideration. Section 6 concludes and makes suggestions for further research.

2 EU ETS and CO_2 trading

2.1 EU ETS

The EU ETS is the key tool of the European Commission to reduce the emissions of GHGs and to comply with the commitments made under the Kyoto protocol. The system is designed using the SO_2 market in the USA as a blueprint and entered into force in 2005 through EU Directive 2003/87/EC. Since then there have been three trading periods. The first trading period, Phase I, lasted from 2005 until 2007 and served as a pilot period to test the market infrastructure. In Phase I the EUAs were freely distributed to the emitting installations. However, the liquidity in the market was low and due to oversupply and the fact that the allowances lost their value at the end of the trading period, prices collapsed towards the end of the trading period. Phase II, which lasted from 2008 until 2012, was the first Kyoto commitment period. Since Phase II banking and borrowing of allowances between years and trading periods is allowed, which reduces the risk of prices to collapse towards the end of the trading period (European Commission, 2012). Both in Phases I and II the allowances were distributed by the principle of grandfathering, i.e. the number of allowances a firm received were relative to the historical emission levels of its installations. The drawback of grandfathering is that it gives rents to existing firms and erects entrance barriers to new firms (Lutz et al., 2013). Therefore, in the current Phase III, which runs from 2013 until 2020, free allocation of EUAs is replaced by auctioning.

Replacing command-and-control regulations to control emissions, the EU ETS is a cap-and-trade system, which means that the regulator, the European Commission, fixes the total amount of emissions and allowances issued in a period. If a firm's emissions exceed the allocated volume of allowances, they can either buy allowances on the market or take abatement measures. Similarly, surplus allowances can be sold. In this way, the right to emit CO_2 becomes a tradable asset. The advantage of cap-and-trade system is that the marginal abatement costs are equalised among the firms, independent of the initial allocation of allowances (Hintermann, 2010). Each year on April 30 firms have to surrender the number of allowances corresponding to the emissions of the previous year. If they fail to do so, the firms have to pay a penalty, 40 EUR and 100 EUR per ton CO_2 emitted in Phases I and II respectively, and to surrender the lacking allowances next year.

The EU ETS created a new market for CO_2 allowances and is now the world's largest carbon market, covering more than 11,000 installations in several sectors. Currently the system covers amongst others power plants, coke ovens, iron and steel factories, and factories producing cement, glass, lime, bricks, ceramics, pulp and paper (European Commission, 2012). The energy sector accounts for roughly half the emissions under the scheme. About half of the total CO_2 emissions in the EU are currently regulated by the EU ETS, while the number of installations included is still growing. Several types of transactions and derivatives have evolved. While in Phase I EUAs were mainly traded in over-the-counter (OTC) transactions, in Phase II they were traded bilaterally, in OTC transactions and on exchanges (Hintermann, 2010). There are spot, future, forward and option markets for EUAs. The spot contracts are traded on several exchanges, amongst others on Bluenext, Climex, European Energy Exchange, Green Exchange, Intercontinental exchange and Nord Pool (European Commission, 2012) of which Bluenext is the largest exchange, covering about 70 per cent of total spot market transactions. Table 2.1 presents the total trade volume on these exchanges, which shows a steady growth both in volume and in traded value (World Bank, 2012).

Year	Number of EUA (in bn.)	Traded value (in USD bn.)
2005	0.3	7.9
2006	1.1	24.4
2007	2.1	49.1
2008	3.1	100.5
2009	6.3	118.5
2010	6.8	133.6
2011	7.9	147.8

Table 2.1: Total spot market trade volumes of EUAs on the six largest exchanges

2.2 Characteristics of EUAs

The characteristics of EUAs and the market setting outlined in this section help to understand the price determinants. Benz & Trück (2006) argue that EUAs are a new type of asset, having different characteristics than traditional stocks or commodities. They argue that EUAs should be considered a factor of production because the right to emit is essential for production. The prices of EUAs are, unlike the prices of stocks, which are determined by expected profits, based on expected market scarcity. It is important to point out that total supply on the market is fixed by the regulator, and that firms can influence their own demand by taking abatement measures. Furthermore, as banking was not allowed between Phase I and II, the EUAs lost their value at the end of the trading period. In addition the market for EUAs is an artificial market created by the EU Directive and thus sensitive to regulatory and policy changes with a potential to influence short-term demand and supply. Finally, during Phase I and II the allowances were distributed free of charge.

The price dynamics of EUAs are governed by shocks, as they depend on factors such as weather, fuel prices and economic growth, which are hard to forecast. Furthermore, the supply and demand is influenced by policy changes, which cannot be forecasted precisely and create unexpected shocks. The European Commission publishes every year a report about the verified emissions under the EU ETS. This is an import signal about the demand side of the market and may create shocks, too. These particularities of the price dynamics should be incorporated in an adequate allowance pricing model.

3 Literature review

Since the creation of the EU ETS there has been an increasing number of studies addressing the modelling of EUA prices. The largest part of the literature concentrates on the determinants and drivers of EUA prices. Furthermore, several studies address the linkage between spot and futures markets for EUAs. The number of studies applying econometric models to EUA data is limited.

Mansanet-Bataller et al. (2007) were one of the the first to analyse the determinants of EUA prices by using empirical data. Using spot and future prices from Phase I, they find evidence for a linkage between fossil fuel prices and EUA price levels. Alberola et al. (2008b) confirm this result and demonstrate that additionally extreme weather events have a strong impact on prices. In Alberola et al. (2007) and Alberola et al. (2008a) they find that additionally the level of economic activity in the main sectors covered by the EU ETS, the energy, steel producing and paper and pulp producing sectors, influence the price process. Moreover, they find two structural breaks in the price process in April and October 2006, when respectively the European Commission published the verified emissions and the announcement of stricter allocations during Phase II. Chevallier (2009) investigates the influence of macroeconomic risk factors on EUA prices but only finds a weak impact. However, Conrad et al. (2010) find clear evidence for the impact of macroeconomic activity as well as for the influence of shocks caused by regulatory information. Finally, Hintermann (2010) derives a structural model for the allowance prices under the assumption of efficient markets and describes how marginal abatement costs influence EUA prices. In conclusion there is clear evidence in the literature for the impact of energy prices, extreme weather events, and economic activity on allowance prices. However, the relation between the allowance prices and these price fundamentals depends on the sample and period considered and changes over time. The previous studies all investigate only data from short time periods, mainly from Phase I. Several authors find structural breaks in price series of EUAs. Alberola et al. (2008b) argue that regulatory changes cause these breaks, whereas Chevallier (2009) sees changes in expectations as the main reason for them. The presence of such breaks complicates the estimation of models for long-term relationships between prices and their fundamentals and calls for endogenising these breaks into the models, such as in regime switching models.

Another strand of research concentrates on the relationship between the spot and futures market for EUAs. Trück et al. (2012) find that the EUA market was in backwardation during Phase I, whereas during Phase II the market moved from backwardation to contango. They apply dynamic semiparametric factor models for modelling the relationship between spot and futures market. Chevallier (2012) applies two nonlinear cointegration models, a VECM with structural shift and a threshold cointegration model, to the EUA spot and futures market. He observes that the returns of spot and futures prices correct the deviations to the long-term equilibrium, with the futures price taking the lead.

Despite the growing importance of carbon finance, few studies have focussed on the stochastic properties of daily EUA spot prices and the application of models from financial econometrics to EUA data. Exceptions are the studies of Paolella & Taschini (2008), Seifert et al. (2008), Daskalakis et al. (2009) and Benz & Trück (2009) which focus on the stochastic properties of daily price data and provide amongst other things evidence for conditional heteroskedasticity. Paolella & Taschini (2008) address the unconditional tail behaviour and heteroskedasticity in the price series by applying mixed GARCH models. However, there findings are only valid for the specific period at the end of Phase I. Seifert et al. (2008) use a stochastic equilibrium model to analyse the dynamics of EUA spot prices. Their main conclusion is that a EAU pricing model should have a time- and price-dependent volatility structure. Daskalakis et al. (2009) model the effects of abolishing banking on futures prices during Phase I and develop a framework for pricing and hedging of intra-phase and inter-phase futures and options on futures. Benz & Trück (2009) use Markov switching and GARCH models for stochastic modelling of the EUA spot prices in Phase I. They find strong support for the use of both types of models to model the characteristics of the series, such as different price phases, volatility clustering, skewness and excess kurtosis. The studies addressing the stochastic proporties of EUA prices are limited to data from Phase I. Due to the peculiarities of the price process in in Phase I as described before, the results are possibly not generalisable to Phase II.

Finally, there is literature on other emission allowance programs, notably on the SO_2 permit trading system in the United States of America. This program has already been in place since 1992. However, the findings relating to the SO_2 market have little relevance for modelling the CO_2 prices in the EU, due to the different market structure and commodity.

4 Methodology

In this section we present the models which we estimate to model the log returns of the EUA spot prices and the log likelihood functions we use for estimation. Besides four regime switching models, we also fit a normal distribution to the data and estimate autoregressive (AR), GARCH and AR-GARCH models in order to have a benchmark for comparing the performance of the regime switching models. Furthermore, we describe the numerical optimisation algorithm used for optimising the log likelihood functions and present measures for comparing in-sample fit and out-of-sample performance of the models.

4.1 Models

4.1.1 AR

AR models can capture the time varying mean of a stochastic process. An AR(p) process with lag order p is defined as

$$y_t = c + \sum_{k=1}^{p} \phi_k y_{t-k} + \varepsilon_t \tag{1}$$

with ε_t a sequence of i.i.d. random variables with mean 0 and variance σ^2 (Hamilton, 1994; Tsay, 2010). We also impose conditions on ϕ_k for $k \in \{1, \ldots, p\}$ to ensure stationarity. When we assume that the error terms are normally distributed, i.e. $\varepsilon_t \overset{i.i.d.}{\sim} \mathrm{N}(0, \sigma^2)$, the vector of parameters to estimate is $\theta_{AR(p)} = (c, \phi_1, ..., \phi_p, \sigma^2)'$ and the maximum likelihood estimator is defined as maximising the log likelihood function as defined in Hamilton (1994). The log likelihood function is only evaluated for the observations $y_{p+1}, \ldots, y_T$, as we need the first p observations for starting the AR process.

4.1.2 GARCH

The AR models as presented in the previous subsection assume homoskedasticity of the error terms. The autoregressive conditional heteroskedastic (ARCH) model of Engle (1982) was the first model to successfully provide a systemic framework to address the issue of heteroskedasticity in time series. The basic idea behind the ARCH model is that the error terms are serially uncorrelated but contain higher-order dependence and can be modelled as a quadratic function of the past error terms (Teräsvirta, 2006). In practice the ARCH model needs many lags to describe the volatility process. In order to avoid this, Bollerslev (1986) proposed a generalization of the ARCH model, the generalised

ARCH (GARCH) model, which includes the own lags of the conditional variance into the ARCH model. y_t follows a GARCH(p,q) model if

$$y_t = \sigma_t \varepsilon_t \tag{2}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{3}$$

where ε_t is a sequence of i.i.d. random variables with mean 0 and variance 1. In this case, the parameter vector to be estimated is $\theta_{GARCH} = (\alpha_0, \dots, \alpha_p, \beta_1, \dots, \beta_q)'$. Usually a GARCH(1,1) process suffices to capture the conditional heteroskedasticity in the series, so the parameter vector reduces to $\theta_{GARCH(1,1)} = (\alpha_1, \beta_1)'$. In order to ensure stationarity and a strictly positive conditional variance, the coefficients have to satisfy $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$ and $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ (Tsay, 2010).

Assuming that the innovations are identically and independently distributed and drawn from a standard normal distribution, i.e. $\varepsilon_t \overset{i.i.d.}{\sim} \mathrm{N}(0,1)$, the conditional maximum likelihood estimator is defined as maximising the following conditional log-likelihood function conditional on its initial values

$$\begin{aligned} &\ell_{GARCH}\left(y_1, y_2, \dots, y_T | \sigma_0^2, \varepsilon_0^2; \theta_{GARCH}\right) \\ &\quad = -\frac{1}{2} \sum_{t=1}^{T} \left(\log(2\pi) + \log(\sigma_t^2) + \frac{\varepsilon_t^2}{\sigma_t^2} \right) \end{aligned} \tag{4}$$

where T is the number of observations and σ_t^2 as defined in Equation 3 (Zivot, 2008). We use the conditional log likelihood, as the unconditional one is not known in closed form. The log likelihood is conditional on the initial values for σ_t^2 and ε_t^2. We use the empirical variance of y_t to initialise the process as proposed by Zivot (2008).

The GARCH model captures the existence of volatility clustering in a more parsimonious way than the ARCH model. In fact a GARCH(p,q) model can be described as a ARCH(∞) model (Teräsvirta, 2006). Furthermore, it can be shown that the tails of a GARCH model with normally distributed error terms are heavier than those of the normal distribution(Tsay, 2010). Heavy tails are often observed in financial time series The unconditional variance $\bar{\sigma}^2$ of a GARCH model is constant and is equal to

$$\bar{\sigma}^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j} \tag{5}$$

The GARCH variance equation can be combined with a specification of the mean. In this paper we use a GARCH process with unconditional mean c. In this case, Equation 2 is replaced by

$$y_t = c + \sigma_t \varepsilon_t \tag{6}$$

with σ_t and ε_t as in Equation 2. We also consider a specification for a conditional mean, an AR(p) process. This replaces Equation 2 with

$$y_t = c + \sum_{k=1}^{p} \phi_k y_{t-k} + \sigma_t \varepsilon_t \tag{7}$$

with σ_t and ε_t as in Equation 2 and certain conditions on ϕ_k for $k \in \{1, \ldots, p\}$ to ensure stationarity. The estimation of both models is equivalent to that of the GARCH model.

4.1.3 Regime switching models

The AR and GARCH models presented in the previous subsections are not able to capture non-linear dynamic patterns in the time series such as breaks or asymmetry. A way to model this non-linearity is the use of regime switching models. The most popular regime-switching model is the Markov regime switching model as proposed by Hamilton (1989). It is an improvement of the random switching model proposed by Quandt (1972) in which the switching is independent over time. Furthermore it performs better than structural change models, because in the latter changes are only modelled as a reaction to identifiable exogenous changes. The Markov regime switching model allows for frequent changes at random points in time, because the regime switching process is governed by a first order Markov chain.

A regime switching model divides the time series into different phases and specifies for each phase a different underlying stochastic process. The phases are also called regimes or states. In this thesis we consider the Markov regime switching model proposed by Hamilton (1989) in which the state variable s_t, which denotes in which state the model is at time t, is a latent, unobservable variable. We restrict ourselves in this paper to models with two states, so that the state space is $\mathcal{S} = \{1, 2\}$. The state at time t is then a realisation of a two-state homogeneous first order Markov chain and is described by the transition probabilities p_{jj} for $j \in \mathcal{S}$, the probability of being in the same state as in the previous period:

$$p_{jj} = \Pr(s_t = j | s_{t-1} = j) \tag{8}$$

Because $p_{ji} = 1 - p_{jj}$ we obtain the transition matrix $\mathbf{P}$

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix} \tag{9}$$

with $p_{ij} = \Pr(s_t = j | s_{t-1} = i)$ denoting the probability of going from state i to state j. Due to the Markov property the current state depends only on the most recent state.
Hamilton (1989) used the Markov regime switching model focussing on the mean behaviour of the variables, but the stochastic process in state j, y_j can also be specified by other models, e.g. conditional variance models.

There are two types of uncertainty when estimating Markov switching models, the unobservable state s_t the stochastic process is in at time t and the population parameters θ_j specifying the process in state j. Inference on the latent state variable can only be made through the observations of y_t as s_t is not observable. To estimate the model we use an iterative algorithm to calculate the conditional log likelihood function as described by Hamilton (1995). The conditional probability that the process is in state j at time t is

$$\xi_{jt} = \mathrm{P}(s_t = j|\Omega_t; \theta) \tag{10}$$

for $j \in \mathcal{S}$, where $\Omega_t = \{y_t, y_{t-1}, ...y_1\}$ are the observations until time t and θ is the parameter vector with the parameters specifying the stochastic process in both states and the transition probabilities. By construction, $\sum_{j=1}^{2} \xi_{jt} = 1$. The inference on the state probabilities ξ_{jt} is performed iteratively by evaluating the density η_{jt} under both regimes

$$\eta_{jt} = g_j(y_t|s_t = j, \Omega_{t-1}; \theta) \tag{11}$$

where g_j is the density function of the process in state j, which depends on the specification of the model and the distribution of the error term. Knowing $\xi_{i,t-1}$ the conditional density of observation y_t is

$$f(y_t|\Omega_{t-1}; \theta) = \sum_{i=1}^{2} \sum_{j=1}^{2} p_{ij} \xi_{i,t-1} \eta_{jt} \tag{12}$$

and the probability to be in state j at time t is

$$\xi_{jt} = \frac{\sum_{i=1}^{2} p_{ij} \xi_{i,t-1} \eta_{jt}}{f(y_t|\Omega_{t-1}; \theta)} \tag{13}$$

This yields the conditional log likelihood of the observed data

$$\ell_{MS}(y_1, y_2, \ldots, y_T|y_0; \theta) = \sum_{t=1}^{T} \ln f(y_t|\Omega_{t-1}; \theta) \tag{14}$$

The maximum likelihood estimator is defined by maximising Equation 14 w.r.t. θ. For the initialisation of ξ we use the approach as suggested by Hamilton (1995) and chose $\xi_{10} = \xi_{20} = \frac{1}{2}$.

4.1.4 GARCH Markov regime switching

The specification of the density function g_j in Equation 11 is straightforward when using a normal distribution or an AR model in the states. In case of a GARCH specification for the conditional variance we encounter a problem with the specification of the volatility. Due to the autoregressive structure of the variance, its specification is path-dependent, it depends on all the preceding unobserved state variables. Hamilton & Susmel (1994) and Cai (1994) were the first to explore Markov regime switching models with ARCH specifications in the states. However, the ARCH effects do not have the problem of path dependency.

The path dependency in the GARCH model makes evaluation of the log likelihood function intractable, as the number of paths grows exponentially with the number of observations. Gray (1996) and Klaassen (2002) made simplifications to the GARCH model to avoid the problem of path dependency and make log likelihood estimation possible. Klaassen uses a first-order recursive procedure for the variance specification.

The variance of y_t in state j evaluated at time $t-1$ is described by Klaassen (2002) as

$$\begin{aligned} \mathrm{Var}_{t-1}(y_t|s_t = j) &= \mathrm{Var}_{t-1}(\varepsilon_t|s_t = j) \\ &= \alpha_{0j} + \alpha_{1j}\varepsilon_{t-1} + \beta_{1j}\,\mathrm{E}_{t-1}\left[\mathrm{Var}_{t-2}(\varepsilon_{t-1}|s_{t-1})\right] \end{aligned} \tag{15}$$

This variance specification integrates out the path dependence by using the law of iterated expectations and has has only a first-order recursive structure.

In this paper we use the specification of the MS-GARCH model as proposed by Klaassen (2002). The advantages of Klaassen's approach are that it allows for recursive estimation of the log likelihood function and for recursive forecasting. Moreover, the regime switching GARCH model solves the problem of volatility persistence encountered in simple GARCH models.

4.2 Maximum likelihood estimation

This section presents the estimation procedure for the log likelihood estimators of the parameter vectors presented in the previous section. Especially in the case of the Markov switching models there are no analytical solutions. Therefore we use a numerical optimisation algorithm in order to estimate the models.

4.2.1 Numerical optimisation

There are several numerical optimisation algorithms available, such as the Newton-Raphson method or Fisher's scoring algorithms. Unfortunately in case of Markov switching models the performance of these algorithms depends on the starting values of the parameters, as they often find only local extrema. Therefore we use the Differential Evolution (DE) algorithm, which does not require the specification of starting values, but is computationally more intensive. The DE algorithm is a genetic algorithm, working with populations and applying crossover, mutation and selection as in biology, which was developed by Storn and Price in the 1990s (Price et al., 2005). DE makes use of artithmetic instead of logical operations and works particularly well to find the global optimum of a real-valued function of real-valued parameters. We use the R package DEoptim, which is developed by Ardia & Mullen (2009).

4.3 Model comparison

In order to evaluate the performance of the different models, we present here several model selection criteria both for the in-sample fit and out-of-sample forecasting performance.

4.3.1 In-sample

The most natural way to compare the goodness-of-fit of the models examined is the value of the log likelihood function. We can compare these values, because all models have the same underlying distribution of the error terms and we use the same sample. The log likelihood of the MS models are naturally higher, due to the increased number of parameters. In order to account for the increased number of parameter we use an information criteria, which introduces a penalty term for the less parsimonious models. The Akaike information criterion (AIC) Akaike (1973) is defined as follows

$$\mathrm{AIC} = -2\ell + 2k \tag{16}$$

$$\tag{17}$$

where ℓ is the value of the estimated log likelihood function and k the number of parameters in the model.

4.3.2 Out-of-sample

To compare the point forecasts of the different models we use the mean absolute error (MAE) and mean squared error (MSE). The MAE and MSE compare the actual value and the forecasted value and are respectively defined as

$$\mathrm{MAE} = \frac{1}{h} \sum_{t=T+1}^{T+h} |\widehat{y}_t - y_t| \tag{18}$$

$$\mathrm{MSE} = \frac{1}{h} \sum_{t=T+1}^{T+h} (\widehat{y}_t - y_t)^2 \tag{19}$$

where $\widehat{y}_t$ is the point forecast for time t, y_t is the true observed value and h is the forecasting horizon.

In order to evaluate the performance of the density forecasts we follow the same approach as Benz & Trück (2009) and perform a distributional test as described in Diebold et al. (1998). This approach is better than the comparison of confidence intervals, as this depends on the choice of the confidence level. Assuming a normal distribution, the forecasted distribution of y_{t+1} is

$$y_{t+1} \sim \mathrm{N}\left(\widehat{\mu}, \widehat{\sigma}^2\right) \tag{20}$$

where $\widehat{\mu}$ is the point forecast and $\widehat{\sigma}^2$ the forecasted variance. If this is the correct distribution with forecasted density function $\hat{f}(y_{t+1})$ and distribution function $\hat{F}(y_{t+1})$, then Rosenblatt (1952) shows that $\hat{F}(y_{t+1})$ is uniformly distributed on the interval $[0, 1]$. The density forecast can now be evaluated by performing a distributional test for uniformity of $\hat{F}(y_{t+1})$. We chose to use the Kolmogorov-Smirnov test.

5 Empirical analysis

5.1 Data

For our empirical analysis we use daily spot market prices from Bluenext in Paris as this is the most liquid market place for spot contracts. The price is for one EUA, which gives the right to emit one ton of CO_2. We utilise data from February 26, 2008 until November 28, 2012, which covers with 1,183 daily observations almost the whole of Phase II. For calibration of the models, we use the data from the period February 26, 2008 until December 30, 2010. The data from January 3, 2011 until November 26, 2012 is used for out-of-sample evaluation of the models. The data was retrieved from Bloomberg with ticker PNXCSPT2. We perform our analysis on log returns of the prices, which are defined as

$$y_t = \log\left(\frac{p_t}{p_{t-1}}\right) \tag{21}$$

where p_t is the daily closing price on the spot market at time t. We use log returns for our analysis in order to obtain well-behaved error terms. Figure 5.1 presents a plot of the daily EUA prices and Figure 5.2 plots the daily log returns. Figure 5.1 shows that the prices in phase II are, contrary to the prices in Phase I, always positive in the period under consideration and have a minimum at 6.04 EUR. The plot of the prices shows as well a decrease of the prices in 2009 and 2011, which corresponds to the effect of the economic crises in both periods. The plot in Figure 5.2 clearly shows volatility clustering and heteroskedasticity. Especially in periods when the price decreases volatility seems to be higher. This can be explained by the fact that the supply of EUAs is inelastic. When the demand decreases due to an external shock, there might rise doubt about the overall shortage of certificates on the market. In case of oversupply the EUAs could become worthless. We will confirm this observation when interpreting the regimes in the regime switching model. Furthermore, we see an increase of volatility in the log returns between February and April of each year. This can be explained by the double bookkeeping in this period. The emitting companies received the allowances for the current year in February and had to surrender the allowances for the previous year at April 30.

Table 5.1 presents the descriptive statistics of both the spot prices and the log returns for the complete time series, the in-sample and out-of-sample period. The log returns are not significantly different from zero. Both the prices and the log returns show excess kurtosis, which mean that the data is heavy-tailed. The prices are positively skewed, whereas the log returns are little skewed. The data is not normally distributed. The observed characteristics of the log return series justify the investigation into MS-GARCH models.

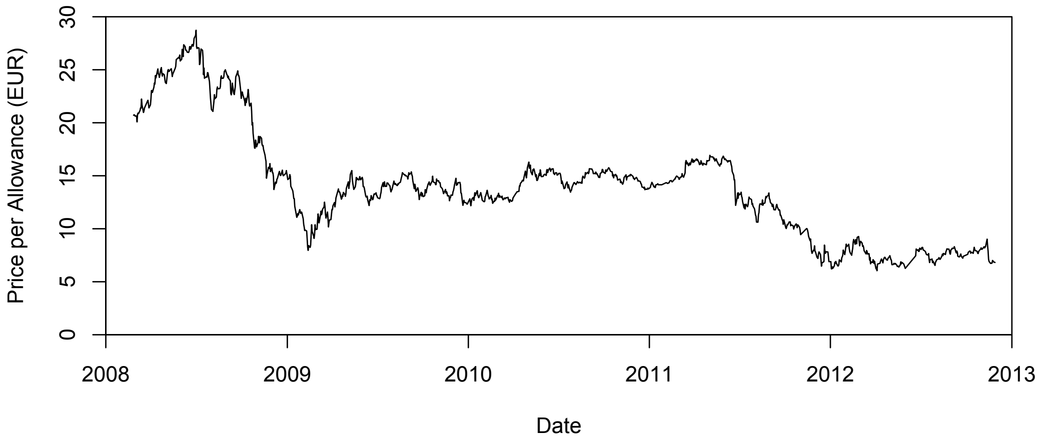

Figure 5.1: Daily EUA prices from February 26, 2008 until November 28, 2012

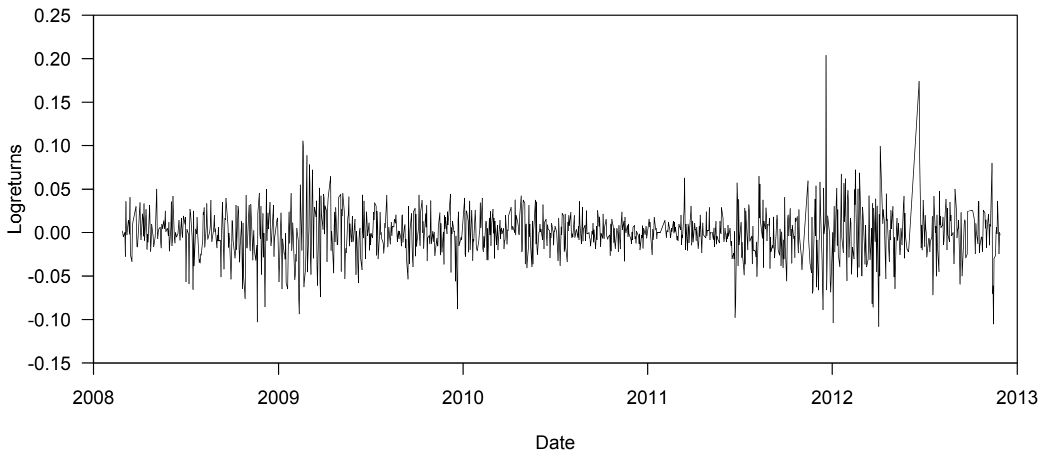

Figure 5.2: Daily EUA logreturns from February 26, 2008 until November 28, 2012

period	N	Mean	Min	Max	Std Dev	Skew	Kurt
			Prices				
2008-2012	1183	14.016	6.040	28.730	5.071	0.76	3.32
2008-2010	725	16.273	7.960	28.730	4.581	1.09	2.99
2011-2012	458	10.433	6.040	16.930	3.505	0.61	8.44
			Log returns				
2008-2012	1182	-0.0009	-0.1081	0.2038	0.0276	0.03	8.03
2008-2010	724	-0.0006	-0.1029	0.1055	0.0244	-0.20	5.02
2011-2012	458	-0.0015	-0.1081	0.2038	0.0320	0.61	8.84

Table 5.1: Descriptive statistics for daily prices and daily log returns

5.2 Estimation results

In this section we present the results of estimating the models on the log returns in the in-sample period. In order to test out-of-sample performance we then forecast the log-returns for the out-of-sample period and finally compare the performance of the models. First we perform stationarity tests on the data.

5.2.1 Stationarity testing

Our models depend on the stationarity assumption of the time series. Therefore we apply both the Augmented Dickey Fuller (ADF) test proposed by Dickey & Fuller (1981) and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test proposed by Kwiatkowski et al. (1990) in order to evaluate the presence of a unit root in the log return series. The ADF test tests the null hypothesis of a unit root against the alternative of a stationary process. The test statistic of the ADF test has a known distribution. We chose the lag order with a general to specific approach, starting with a maximum lag order of $(T-1)^{\frac{1}{3}}$. The KPSS test tests the null hypothesis that the time series is stationary around a deterministic trend against the alternative hypothesis of a unit root process. The KPSS test is more conservative than the ADF test. Table 5.2 presents the test statistics, the p-values and the used lag orders for both stationarity tests. The ADF test reject the null hypothesis of a unit root process and the KPSS test accepts the null hypothesis of a stationary process. Both tests come to the same conclusion for all three periods at high significance levels.

period	test statistic	p-value	lags
	ADF		
2008-2012	-7.437	<0.01	22
2008-2010	-5.321	<0.01	20
2011-2012	-4.879	<0.01	17
	KPSS		
2008-2012	0.069	>0.1	7
2008-2010	0.108	>0.1	6
2011-2012	0.071	>0.1	4

Table 5.2: Results of the ADF and KPSS tests for stationarity

5.2.2 Normal distribution and AR model

In this subsection we present the results of fitting a normal distribution and estimating an AR model to the log returns. The parameter estimates for the fitted normal distribution (i.i.d. Normal) and the AR(4) model are presented in Table 5.3. We first test the autocorrelation structure in the log returns. The autocorrelation function (ACF) in the upper panel of Figure 5.3 shows the presence of autocorrelation, as the spikes at lag orders 1, 2 and 4 are significant. In order to determine the lag order of the AR process we follow the approach of Tsay (2010), which uses the sample partial ACF (PACF) and the AIC. The sample PACF of the logreturns in the lower panel of Figure 5.3 shows that a typical pattern of an AR process. Its spikes until lag order 3 are significant, which suggest an AR(3) process. However, according to the AIC, we prefer an AR(4) process. As the spike at lag 4 of the PACF is also large, we estimate an AR(4) process.

The estimated mean of the normal distribution and unconditional mean of the AR(4) model are almost the same. Also the estimated variance is almost the same for both models. This indicates that the additional explanatory power of the AR model is rather limited. The coefficients of the AR process suggest that sign changes in the log returns are rather limited.

5.2.3 GARCH

This subsection presents the parameter estimates of the GARCH and AR-GARCH models. Volatility clustering or GARCH effects in the data can be detected by autocorrelation in the squared or absolute returns in the series or in the residuals of an estimated model for the mean. The latter is in our case the fitted normal distribution or the AR model. The upper panel in Figure 5.4 plots the residuals of the AR(4) model, which we estimated in the previous section and the lower panel shows the ACF of the squared residuals. The upper plot shows a non-constant variance and the lower plot correlation

Parameter	Coefficient	Coefficient
	i.i.d. Normal	AR(4)
μ	-0.0006	–
c	–	-0.0006
ϕ_1	–	0.0988
ϕ_2	–	-0.1391
ϕ_3	–	0.0795
ϕ_4	–	0.0609
σ	0.0244	0.0240
Unconditional expectations		
$\mathrm{E}[y_t]$	-0.0006	-0.0006
σ	0.0244	0.0240

Table 5.3: Parameter estimates of i.i.d. Normal and AR models

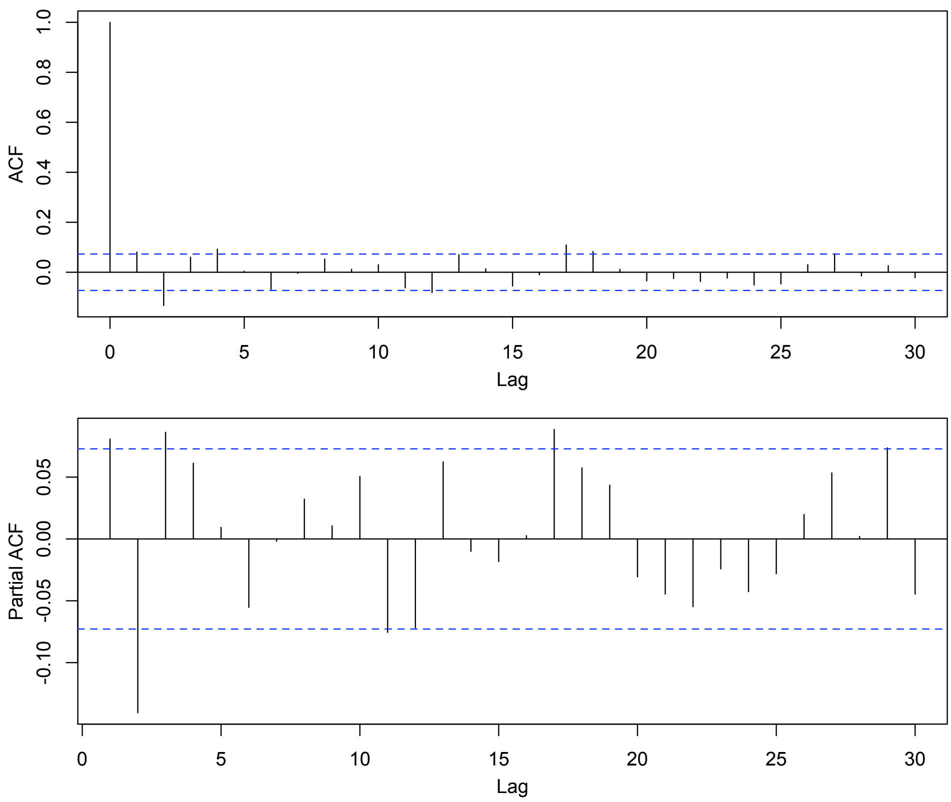

Figure 5.3: ACF (upper panel) and PACF (lower panel) of the log return series from February 26, 2008 until December 30, 2010
Dotted blue lines give approximate 95% confidence bands

in the residuals, which both are indicators for GARCH effects. To test this intuition we use the Ljung-Box or modified Q-statistic proposed by Box & Pierce (1978)

$$\mathrm{MQ}(p) = T(T+2)\sum_{j=1}^{p}\frac{\hat{\rho}_j^2}{T-j} \tag{22}$$

where $\hat{\rho}_j$ denotes the j-lag sample autocorrelation and T the sample size. Under the null hypothesis that the data are white noise the test statistic $\mathrm{MQ}(p)$ has an asymptotic χ^2-squared distribution with p degrees of freedom. Table 5.4 shows the results of this test on the squared residuals of the estimated AR(4) model. The test rejects the null hypothesis of white noise and confirms the presence of GARCH effects in the data.

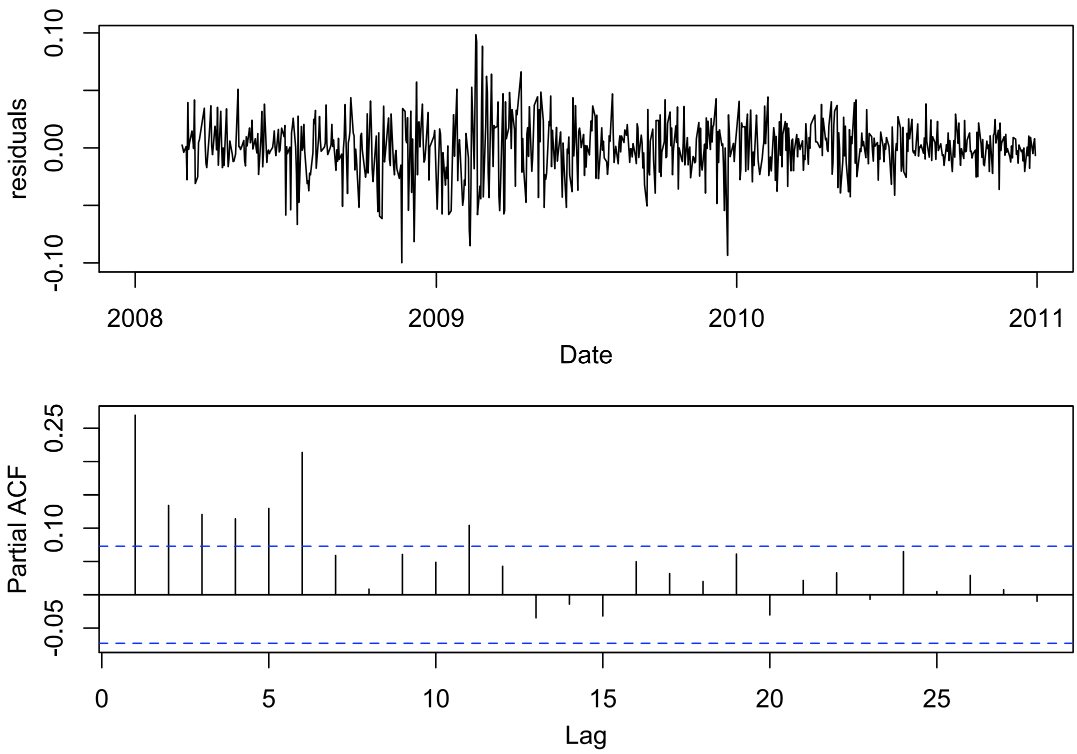

Figure 5.4: Residuals of the AR(4) model (upper panel) and ACF of the squared residuals of the AR(4) model (lower panel)

test statistic	p-value
61.793	3.775e-15

Table 5.4: Results of Ljung-Box test for GARCH effects in squared residuals of the AR(4) model

Table 5.5 shows the parameter estimates of the GARCH(1,1) and the AR(4)-GARCH(1,1) models. All parameters are significant in both models, except for α_0. We also estimated higher order GARCH models, however the coefficients were not significant. For the AR-GARCH model we choose the same lag order as in the AR model estimated before. For both GARCH models the sum of the parameters α_1 and β_1 is close to one and α_0 is small which indicates a high persistence of the volatility and a slow reversion to the mean. This is a well-known drawback of GARCH modelling of financial time series. In both models we observe an unconditional mean smaller than the mean of the series. The unconditional standard deviation is close to the empirical standard deviation of the series.

Parameter	GARCH(1,1)	AR(4)-GARCH(1,1)
Mean equation		
c	-0.0003	-0.0002
ϕ_1	–	0.0031
ϕ_2	–	-0.0696
ϕ_3	–	0.0550
ϕ_4	–	0.0199
Variance equation		
α_0	0.0000	0.0000
α_1	0.0726	0.0697
β_1	0.9199	0.9214
Unconditional expectations		
$E[y_t]$	-0.0003	-0.0002
$E[\sigma_t]$	0.0239	0.0247

Table 5.5: Parameter estimates of GARCH(1,1) and AR(4)-GARCH(1,1) models

5.2.4 Gaussian and AR Markov regime switching

Table 5.6 presents the in-sample parameter estimates of the Markov regime switching models with a fitted normal density (MS-i.i.d. Normal) and an AR(4) process (MS-AR(4)) in the regimes. In both models one state is characterised by low volatility and a positive mean ('low') and the other state is characterised by high volatility and a negative mean ('high'). The 'low' state can be interpreted as the base or normal state and the 'high' state as a period of uncertainty. This uncertainty is a result of the design of the system. An unexpected event, such as a drop in economic activity or regulatory announcement could reduce CO_2 production and thus the demand for EUAs and result in a falling price. As the supply side is fixed, there might be uncertainty on the market, whether the demand will be higher than the supply, which hence causes higher volatility.

The variance in the 'high' state is in both models more than four times higher. This allows for sudden changes from low to high volatility by a regime change in the model. These changes are clearly visible in the estimated state probabilities in the upper panel of Figure 5.5. In the upper panel the estimated probability to be in the 'low' regime for the MS-AR(4) model is plotted. In periods of high volatility (identifiable in the lower panel of Figure 5.5) the probability to be in the low regime drops suddenly, which means that the the probability to be in the high regime is very high as these probabilities sum up to 1. At any point in time the model assigns with high probability one of both regimes, which means that the model distinguishes well between the states.

In both models the unconditional probability to be in the 'low' regime is much higher, 65% and 62% for respectively the MS-i.i.d. Normal and MS-AR(4) models. The transition probabilities to stay in the same regime are very high, close to 100% for both regimes in both models. This indicates that regime changes are rather rare. The results for the MS-i.i.d. Normal and MS-AR(4) models are similar.

	MS-i.i.d. Normal		MS-AR(4)	
Regime (i)	1 (low)	2 (high)	1 (low)	2 (high)
μ_1	0.0014	-0.0037	–	–
σ_i	0.0161	0.0336	0.0159	0.0324
c	–	–	0.0017	-0.0033
ϕ_1	–	–	-0.0597	0.1647
ϕ_2	–	–	-0.0662	-0.1947
ϕ_3	–	–	0.0086	0.1116
ϕ_4	–	–	-0.0870	0.1078
Markov estimates				
p_{ii}	0.9864	0.9749	0.9818	0.9698
Unconditional expectations				
$\mathrm{E}[y_{t,i}]$	0.0014	-0.0037	0.0014	-0.0041
$\mathrm{E}[\sigma_{t,i}]$	0.0161	0.0336	0.0159	0.0324
$\mathrm{P}(s_t = i)$	0.6486	0.3514	0.6240	0.3760

Table 5.6: Parameter estimates of Markov switching i.i.d. Normal and AR(4) models

5.2.5 GARCH Markov regime switching

Table 5.7 presents the estimated parameters of the MS-GARCH(1,1) and MS-AR(4)-GARCH(1,1) models. We observe the same 'low' and 'high' states as in the previous MS models. In the MS-AR(4)-GARCH(1,1) model the unconditional standard deviation in the 'high' state is even seven times higher than in the 'low' state. The transition probabilities to stay in the same state are very high for the MS-GARCH(1,1) model, which means that the number of regime switches is limited. For the MS-AR(4)-GARCH(1,1) model, we observe a lower transition probability to stay in the 'high' state, which means

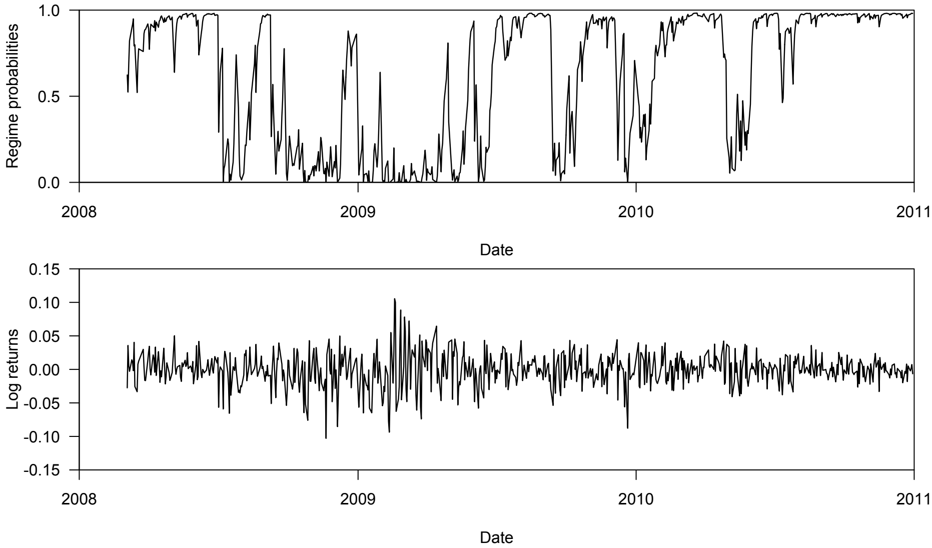

Figure 5.5: Estimated probabilities to be in the 'low' state for MS-AR(4) (upper panel) model and log returns (lower panel)

that this state is less stable. This is also reflected in the unconditional probability to be in the 'high' state, which is only 18%, opposed to 82% for the 'low' state. The MS-AR(4)-GARCH(1,1) model distinguishes well between the regimes. This is shown in Figure 5.6, which is analogous to Figure 5.5. Again the probability to be in the 'low' regime drops in times of high observed volatility. However, the regime selection is not as pronounced as in the case of the MS-AR(4) model.

5.2.6 Comparison

Table 5.8 presents measures for the in-sample goodness-of-fit of our models. It is possible to compare the log likelihood values and information criteria, because we use for all models the same sample and distributional assumption. According to the log likelihood value, the Markov switching models have a better fit than the standard model with the same specification. This result is confirmed by the AIC, which accounts for the parsimony of the models. The best in-sample fit has the MS-AR(4)-GARCH(1,1) model, according to the log likelihood and the AIC. Especially the Markov switching models have many parameters to estimate. The MS-GARCH models perform better than the MS models without a GARCH specification. The GARCH(1,1) and AR(4)-GARCH(1,1) models have a better in-sample fit than the MS models without GARCH specification according to the AIC. This contradicts the findings of Benz & Trück (2009), who found a similar in-sample fit for the GARCH and simple Markov switchig models. Finally, we notice

	MS-GARCH(1,1)		MS-AR(4)-GARCH(1,1)	
Regime (i)	1 (low)	2 (high)	1 (low)	2 (high)
Mean equation				
c	0.0009	-0.0042	0.0011	-0.0090
ϕ_1	–	–	-0.0339	0.3013
ϕ_2	–	–	-0.0637	-0.2108
ϕ_3	–	–	0.0261	0.1965
ϕ_4	–	–	-0.0315	0.2512
Variance equation				
α_0	0.0001	0.0003	0.0000	0.0002
α_1	0.0013	0.1038	0.0078	0.1952
β	0.7166	0.7233	0.8645	0.7510
Markov estimates				
p_{ii}	0.9923	0.9821	0.9740	0.8818
Unconditional expectations				
$\mathrm{E}[y_{t,i}]$	0.0009	-0.0042	0.0010	-0.0218
$\mathrm{E}[\sigma_{t,i}]$	0.0136	0.0409	0.0101	0.0707
$\mathrm{P}(s_t = i)$	0.6988	0.3012	0.8198	0.1802

Table 5.7: Estimates of Markov switching GARCH(1,1) and AR(4)-GARCH(1,1) models

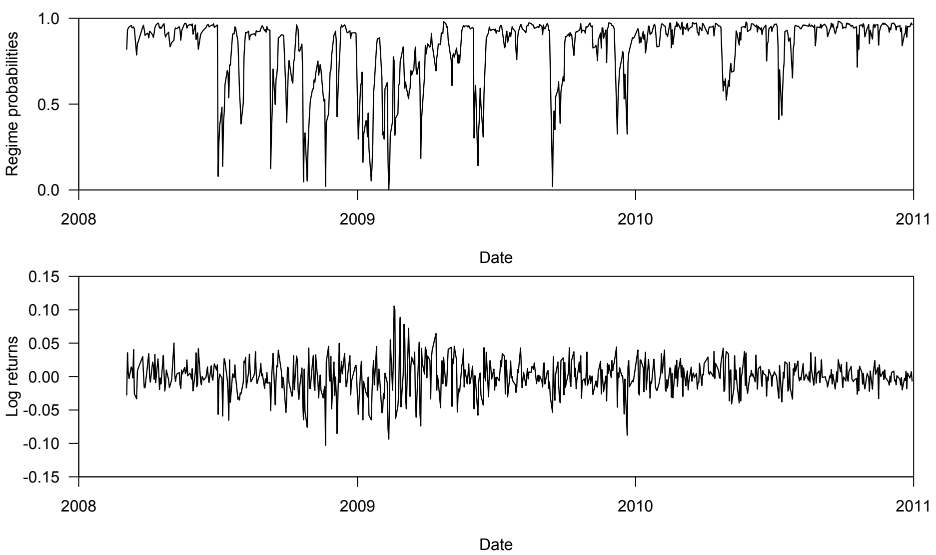

Figure 5.6: Estimated probabilities to be in the 'low' state for MS-AR(4)-GARCH(1,1) model (upper panel) and log returns (lower panel)

that the autoregressive mean specification provides a better sample fit in all the models.

model	number of parameters	log likelihood	AIC
i.i.d. Normal	2	1651.06	-3298.11
AR(4)	6	1673.85	-3335.69
GARCH(1,1)	4	1732.45	-3456.89
AR(4)-GARCH (1,1)	8	1735.33	-3454.67
MS i.i.d.	6	1720.00	-3408.99
MS-AR(4)	14	1732.92	-3437.84
MS-GARCH(1,1)	10	1739.21	-3458.43
MS-AR(4)-GARCH(1,1)	18	1750.94	-3465.87

Table 5.8: Number of parameters, maximum log likelihood value and Akaike information criteria (AIC) for the estimated models

5.3 Forecasting

Forecasting of prices and volatility is important for risk management. Therefore we compare the forecasting performance of the previous eight models by performing out-of-sample forecasts. We make one-day-ahead forecasts for the period from January 3, 2011 until November 26, 2012 and compare these forecasts with the true observed values. We use a recursive window approach in which we reestimate the model every day by using all previous data points since February 26, 2008. In this way the sample size increases when estimating and forecasting later log returns. The reestimation of the parameters is expected to improve the forecasting performance. Besides point forecasts for the log returns, we also focus on density forecasts, as these are often more relevant to risk managers. Also density forecast allow to construct confidence intervals. We evaluate the forecasts by using the techniques described in Section 4.3.2.

The point forecasts are evaluated by calculating the average forecast error. Table 5.9 presents the MAE and MSE for all models. The smallest MAE is observed for the MS-AR(4)-GARCH(1,1) model, which has the second smallest MSE. The performance of the fitted normal distribution is remarkable. However, the differences in the values for MAE and MSE are small. This might be due to the short forecasting horizon. We therefore conclude that the results for the mean forecasting are without substantial differences.

We also made density forecast based on the normality assumption, which allows to forecast confidence intervals. Figures 5.7 and 5.8 plot the forecasted confidence intervals (black), the point forecasts (red) and the true values (blue). We observe smaller confidence intervals for the MS-GARCH models. Especially we see that the problem of volatility persistence is reduced by the MS-GARCH model, when comparing the confi-

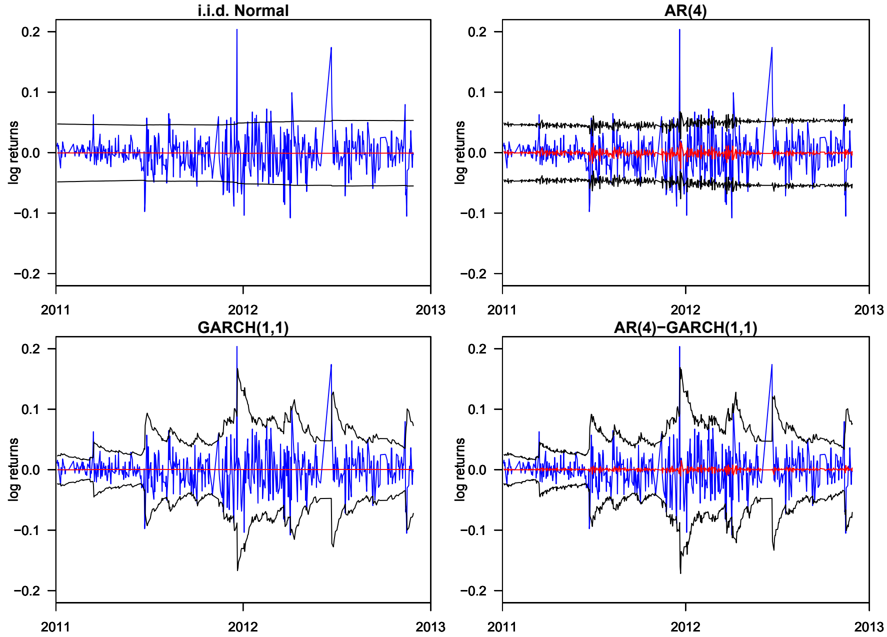

Figure 5.7: Forecasted confidence intervals (black), point forecasts (red) and true values (blue) I

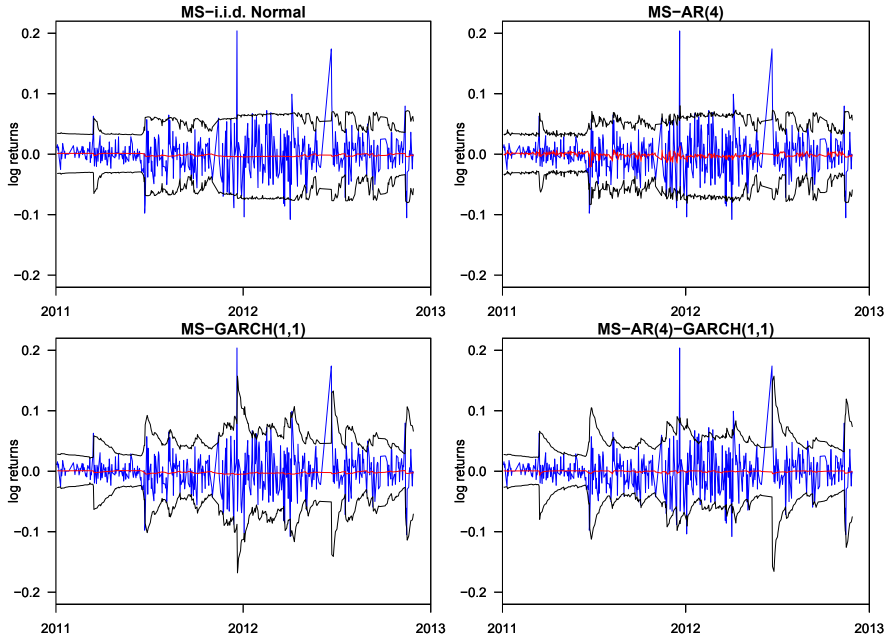

Figure 5.8: Forecasted confidence intervals (black), point forecasts (red) and true values (blue) II

dence intervals of the MS-GARCH models with those of the GARCH models. To test this observation, we use the density test as described in Section 4.3.2. The results of the Kolmogorov-Smirnof test are presented in Table 5.9. The results for the MS models are much better. The best density forecasts are made with MS-GARCH(1,1) model.

model	MAE	MSE	KS	p-value KS
i.i.d. Normal	0.02226	0.0010263	0.4737	<2.2e-16
AR(4)	0.02244	0.0010583	0.0469	0.2657
GARCH(1,1)	0.02230	0.0010282	0.0536	0.1446
AR(4)-GARCH (1,1)	0.02231	0.0010391	0.0501	0.2005
MS i.i.d.	0.02234	0.0010266	0.0367	0.5695
MS-AR(4)	0.02260	0.0010407	0.0346	0.6419
MS-GARCH(1,1)	0.02232	0.0010254	0.0321	0.7314
MS-AR(4)-GARCH(1,1)	0.02229	0.0010268	0.0370	0.5592

Table 5.9: Mean absolute error (MAE) and mean squared error (MSE) for point forecasts and Kolmogorov-Smirnov (KS) test for density forecasts

Our models are based on the assumption of normality of the error terms. Figures 5.9 and 5.10 show kernel density plots of the standardised forecast errors. We see that the standardised forecast errors for the non-MS models seem to have heavier tails than the normal distribtuion. The MS models show almost normally distributed standardised forecast errors. In order to test this intuition we perform both a the Shapiro-Wilk test (Shapiro & Wilk, 1965) and a Kolmogorov-Smirnov test for normality of the standardised forecast errors. The Shapiro-Wilk tests the null hypothesis of normality, which is rejected if the value of the test statistic is close to zero. Values close to 1 support the null hypothesis. The Kolomogorov-Smirnov test is a non-parametric test with the null hypothesis of normality. The test statistic follows the Kolmogorov distribution. The results of both tests are presented in Table 5.10. We do not reject the null hypothesis for any of the Markov switching models. Also for all standard models we do not reject the null hypothesis of normality.

model	SW test statistic	KS test statistic	p-value
i.i.d. Normal	0.935	0.0396	0.4701
AR(4)	0.933	0.0469	0.2657
GARCH(1,1)	0.936	0.0536	0.1446
AR(4)-GARCH (1,1)	0.937	0.0498	0.2067
MS i.i.d.	0.946	0.0367	0.5695
MS-AR(4)	0.949	0.0346	0.6419
MS-GARCH(1,1)	0.936	0.0321	0.7314
MS-AR(4)-GARCH(1,1)	0.941	0.0370	0.5592

Table 5.10: Results of Shapiro-Wilk (SW) and Kolmogorv-Smirnov (KS) tests for normality of the forecast errors

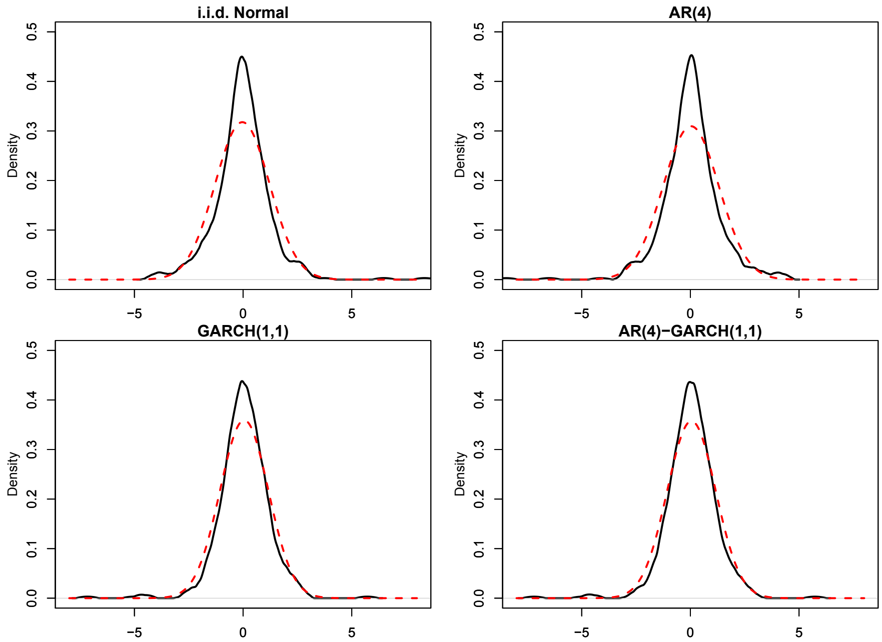

Figure 5.9: Kernel density plots of standardised forecast errors (black solid line) and normal densities (red dashed line)

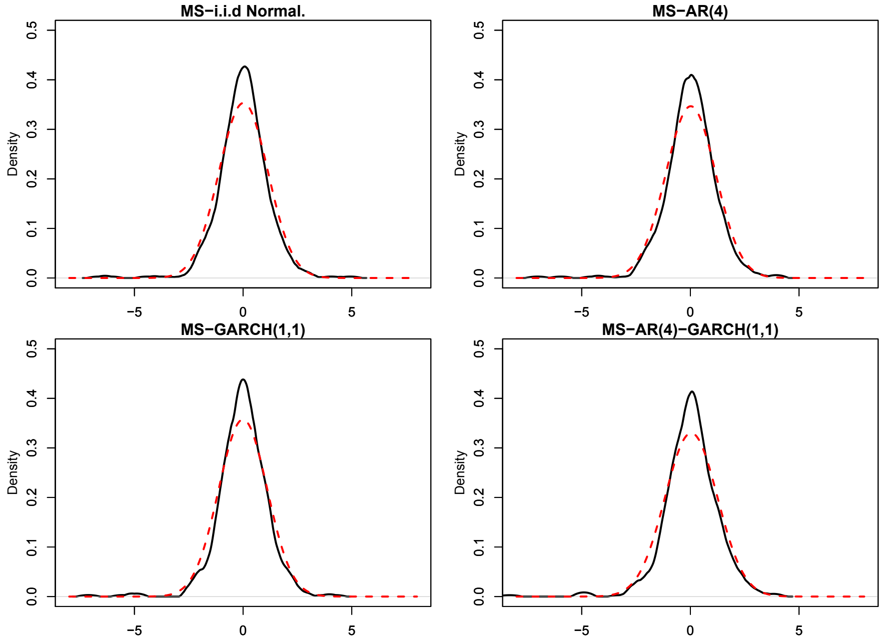

Figure 5.10: Kernel density plots of standardised forecast errors (black solid line) and normal densities (red dashed line)

6 Conclusion

In this paper we studied the short-term spot price behaviour of EUAs, which is of particular importance in the transition of energy markets and for the development of new risk management strategies. Emphasis was given to short-term forecasting of prices and volatility. We took a similar approach as Benz & Trück (2009) but analysed the log returns of Phase II spot market prices and extended the approach by investigating also MS-GARCH models. The application of MS-GARCH models is suggested by different authors and justified by the characteristics of the observed series. These characteristics are, amongst others, volatility clustering, breaks in the volatility process and heavy-tailed distributions.

We estimated eight models to the data: a normal distribution, an AR(4), a GARCH(1,1) and an AR(4)-GARCH(1,1) model, all with and without regime switching. The best in-sample fit to the data is achieved by the MS-AR(4)-GARCH(1,1) model. The results of the MS models are clearly better than those of the non-switching models. The Markov regime switching models estimate two clearly different regimes, a 'low' regime with low volatility and a high mean and a 'high' regime with high volatility and a low mean. The 'low' state can be interpreted as the base or normal state and the 'high' state as a period of uncertainty. This uncertainty is a result of the design of the EU ETS. An unexpected event, such as a drop in economic activity or regulatory announcement could reduce CO_2 production and thus the demand for EUAs and result in a falling price. As the supply side is fixed, there might be uncertainty on the market, whether the demand will be higher than the supply, which hence causes higher volatility. The regime switching models distinguish well between the 'high' and 'low' states.

Concerning point forecasts of the prices, we observe small differences in the MAE and MSE for the different models. The results of the fitted normal distribution are remarkable. We also conducted density forecasts and evaluated their performance by using a distributional test. The best results were achieved by the MS-GARCH(1,1) model. Observing the forecasted confidence intervals, we see that the MS-GARCH models solve the issue of volatility persistence observed in GARCH models. Our results strongly support the use of MS-GARCH models for volatility forecasting for risk management.

So far, we applied only models with normally distributed error terms. Although the modelling of fat tails is partially addressed by GARCH models, we suggested to investigate the use of other heavy-tailed distributions, such as the Student's t-distribtion as suggested by Klaassen (2002) for MS-GARCH models.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. Petrov, & F. Csaki (Eds.) *Second international symposium on information theory*, (pp. 267–281). Budapest: Academiai Kiado.

Alberola, E., Chevallier, J., & Chèze, B. (2007). European carbon prices fundamentals in 2005-2007: the effects of energy markets, temperatures and sectorial production. EconomiX Working Papers 2007-33, University of Paris West - Nanterre la Défense, EconomiX.

Alberola, E., Chevallier, J., & Chèze, B. (2008a). The EU emissions trading scheme : The effects of industrial production and CO_2 emissions on carbon prices. Open Access publications from Université Paris-Dauphine urn:hdl:123456789/4223, Université Paris-Dauphine.

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

Ardia, D., & Mullen, K. (2009). Deoptim: Differential evolution optimization in R. URL `http://CRAN.R-project.org/package=DEoptim`

Benz, E., & Trück, S. (2006). CO_2 emission allowances trading in Europe – Specifying a new class of assets. *Problems and Perspectives in Management*, *4*, 30–40.

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

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Box, G. E. P., & Pierce, D. A. (1978). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, *65*, 1509–1526.

Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business & Economic Statistics*, *12*(3), 309–16.

Chevallier, J. (2009). Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Economics*, *31*(4), 614–625.

Chevallier, J. (2012). Cointegration between carbon spot and futures prices : from linear to nonlinear modeling. Open Access publications from Université Paris-Dauphine urn:hdl:123456789/7936, Université Paris-Dauphine.

Conrad, C., Rittler, D., & Rotfuß, W. (2010). Modeling and explaining the dynamics of European Union allowance prices at high-frequency. ZEW Discussion Papers 10-038, ZEW - Zentrum für Europische Wirtschaftsforschung / Center for European Economic Research.

Daskalakis, G., Psychoyios, D., & Markellos, R. N. (2009). Modeling CO_2 emission allowance prices and derivatives: Evidence from the European trading scheme. *Journal of Banking & Finance*, *33*(7), 1230–1241.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, *49*, 1057–1072.

Diebold, F. X., Gunther, T. A., & Tay, A. S. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review*, *39*, 863–883.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987–1007.

European Commission (2012). The EU emissions trading system (EU ETS). URL `http://ec.europa.eu/clima/policies/ets/index_en.htm`

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, *42*(1), 27–62.

Hamilton, J. (1995). Regime-switching models. URL `dss.ucsd.edu/~jhamilto/palgrav1.pdf`

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*(2), 357–84.

Hamilton, J. D. (1994). *Time series analysis*. SpringerLink : Bücher. Princeton University Press.

Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, *64*, 307–333.

Hintermann, B. (2010). Allowance price drivers in the first phase of the EU ETS. *Journal of Environmental Economics and Management*, *59*(1), 43–56.

Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. *Emperical Economics*, *27*, 363–394.

Kwiatkowski, D., Phillips, P., Schmidt, P., & Shin, Y. (1990). Testing the null hypothesis of stationarity against the alternative of unit root : How sure are we that economic time series have a unit root? Papers 8905, Michigan State - Econometrics and Economic Theory.

Lutz, B. J., Pigorsch, U., & Rotfu, W. (2013). Nonlinearity in cap-and-trade systems: The EUA price and its fundamentals. ZEW Discussion Papers 13-001, ZEW - Zentrum fr Europische Wirtschaftsforschung / Center for European Economic Research.

Mansanet-Bataller, M., Pardo, A., & Valor, E. (2007). CO_2 prices, energy and weather. *The Energy Journal*, *0*(3), 73–92.

Paolella, M. S., & Taschini, L. (2008). An econometric analysis of emission allowance prices. *Journal of Banking and Finance*, *32*(10), 2022–2032.

Price, K., Storn, R., & Lampinen, J. (2005). *Differential Evolution: A practical approach to global optimization*. Natural Computing Series. Springer.

Quandt, R. (1972). A new approach to estimating switching regressions. *Journal of the American Statistical Association*, *67*, 306–310.

Rosenblatt, M. (1952). Remarks on a multivariate transformations. *Annals of Mathematical Statistics*, *23*, 470–472.

Seifert, J., Uhrig-Homburg, M., & Wagner, M. (2008). Dynamic behavior of CO_2 spot prices. *Journal of Environmental Economics and Management*, *56*(2), 180–194.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3-4), 591–611.

Teräsvirta, T. (2006). An introduction to univariate GARCH models. Working Paper Series in Economics and Finance 646, Stockholm School of Economics.

Trück, S., Härdle, W., & Weron, R. (2012). The relationship between spot and futures CO_2 emission allowance prices in the EU-ETS. HSC Research Reports HSC/12/02, Hugo Steinhaus Center, Wroclaw University of Technology.

Tsay, R. S. (2010). *Analysis of financial time series*. Wiley series in probability and statistics. Wiley.

World Bank (2012). *State and trends of the carbon market 2012*. World Bank.

Zivot, E. (2008). Practical issues in the analysis of univariate GARCH models. Working Papers UWEC-2008-03-FC, University of Washington, Department of Economics.

Declaration of authorship

I hereby confirm that I have authored this master thesis independently and without use of others than the indicated sources. Where I have consulted the published work of others, in any form (e.g. ideas, equations, figures, text, tables), this is always explicitly attributed.

Berlin, March 4th, 2013 Thijs Benschop