Forecasting the European Carbon Market

Gary Koop & Lise Tole
Department of Economics
University of Strathclyde

February 2011, revised August 2011

Abstract: In an effort to meet its obligations under the Kyoto Protocol, in 2005 the European Union introduced a cap-and-trade scheme where mandated installations are allocated permits to emit CO$_2$. Financial markets have developed that allow companies to trade these carbon permits. For the EU to achieve reductions in CO$_2$ emissions at a minimum cost, it is necessary that companies make appropriate investments and policymakers design optimal policies. In an effort to clarify the workings of the carbon market, several recent papers have attempted to statistically model it. However, the European carbon market (EU ETS) has many institutional features that potentially impact on daily carbon prices (and associated financial futures). As a consequence, the carbon market has properties that are quite different from conventional financial assets traded in mature markets. In this paper, we use dynamic model averaging (DMA) in order to forecast in this newly-developing market. DMA is a recently-developed statistical method which has three advantages over conventional approaches. First, it allows the coefficients on the predictors in a forecasting model to change over time. Second, it allows for the entire forecasting model to change over time. Third, it surmounts statistical problems which arise from the large number of potential predictors that can explain carbon prices. Our empirical results indicate that there are both important policy and statistical benefits with our approach. Statistically, we present strong evidence that there is substantial turbulence and change in the EU ETS market, and that DMA can model these features and forecast accurately compared to conventional approaches.

Keywords: Bayesian, carbon permit trading, financial markets, state space model, model averaging

---

1We would like to thank Beat Hintermann, two referees and seminar participants at Envecon 2011 for helpful comments. Financial support from the ESRC under grant RES-062-23-2646 is gratefully acknowledged. Both authors are Fellows at the Rimini Centre for Economic Analysis.

2Address for correspondence: Lise Tole, Department of Economics, University of Strathclyde, 130 Rottenrow, Glasgow G4 0GE, U.K. Email: lisetole@googlemail.com.
1 Introduction

The largest carbon market in the world, the European Union Trading Scheme (EU ETS) commenced operation on January 1st, 2005. The scheme caps CO$_2$ emissions, and then distributes allowances to emit to large energy users. One EU allowance (EUA) is equal to one metric tonne of CO$_2$. This "right to pollute" permit can be traded on financial markets. The supply of permits is set by the cap, and the demand for permits depends on the level of CO$_2$ emissions in a given year. In 2009, the total value of the market had grown to €103 billion by the end of the year, with 8.7 billion tonnes of CO$_2$ traded, accounting for over 95% of the total value of carbon traded worldwide (Kossoy and Ambrosi 2010). There has been a growing interest in statistically modelling this increasingly important financial asset.

Like any commodity market, certain fundamentals (referred to as “price drivers”) are expected to play a role in explaining changes in the carbon price. But the exact list of price drivers and the magnitude of their impacts on price is unclear. Moreover, unlike other commodity markets, the European carbon market is fundamentally determined by compliance. The supply of permits is set by the European Commission (EC), which also controls the allocation, bankability and size of the program through the number of participants and permits. These elements of the trading scheme have differed between the first (2005-2007) and the second trading phase (2008-2012). They will also change again, in the third phase (2013-2020), as the cap tightens and the size of the market increases with the extension to more polluters. Uncertainty over a global climate agreement, the size of the future cap and market, the eligibility and number of some types of offsets allowable in the scheme are key factors impacting on the price of carbon now and in the future. While all commodities markets face uncertainties, the changing conditions imposed by compliance are unique to the EU ETS.

Because of the changing complexities of this commodity market, measuring price drivers and forecasting prices raises particular challenges for the statistician seeking to understand the dynamics of this new and evolving market. As with many financial studies (e.g. Avramov 2002), the number of potential variables that may affect prices is large. This can lead to over-parameterization problems (i.e. regression techniques may apparently fit well in-sample, but this may be due to over-fitting, leading to poor forecast performance). In response to this problem, it is increasingly common to use Bayesian model averaging in such cases (see e.g. Avramov, 2002, Ley and Steel, 2009 and 2010 and the references cited therein). Moreover, the marginal effects of predictors can change over time, i.e. parameters can change or structural breaks can occur. Also, the relevant forecasting model may change. These types of changes are especially likely to occur in a market such as the EU ETS, where there are many potential macroeconomic, financial and institutional variables, the impacts of which will vary both within and between phases, with implications for the price of carbon. To address these issues, we use an approach called Dynamic Model Averaging (DMA). DMA can handle the over-fitting problems caused by the presence of
many predictors and allows for both parameter and model change and, thus, seems ideal for forecasting the carbon markets.

We forecast spot and future prices of carbon permits traded between April 2005 and August 2010 using a set of large set of potential predictors which combines those used in a variety of other studies. We find that DMA forecasts of carbon prices in the EU ETS are better than conventional forecasts. Furthermore, DMA gives insight into the range and changing role of various factors driving carbon prices in the different phases of the EU ETS.

The remainder of this paper is organized as follows: Section 2 provides an overview of the EU ETS market since 2005, its main institutional aspects and price drivers. It includes an overview of the empirical literature on carbon trading as it relates to the EU ETS. This section also motivates the choice of variables used in the analysis. Section 3 discusses the data. Section 4 presents the DMA methodology. Section 5 presents the paper’s results. Section 6 concludes. An online appendix, available at http://personal.strath.ac.uk/gary.koop/koop_tole_appendix.pdf provides additional discussion of institutional details, the data and further empirical results.

2 Carbon Prices in the EU ETS

2.1 Institutional Aspects

Under the Kyoto Protocol, the EU has agreed to cut emissions by 8% relative to 1990 levels by the years 2008-12. This may rise to as much as 30% relative to 1990 levels by the end of 2020. The EU ETS is the key institution whereby EU countries aim to meet their Kyoto obligations. Large emitters subject to compliance are allocated an annual number of permits. Currently, over 11,000 energy-intensive installations in the power industry and five major industrial sectors (including oil, iron and steel, cement, glass, and pulp and paper) are included in the scheme. Together these account for nearly half of Europe’s total CO\textsubscript{2} emissions. Each permit grants the right to emit one tonne of CO\textsubscript{2}. Permits can be bought and sold in financial markets. An installation that exceeds its cap can purchase additional permits on the market. An installation that has more permits than emissions can sell its excess permits or bank them for a limited period. At the end of each year, companies must surrender a number of permits equal to the amount of their total emissions for the year. Each year the EC discloses information on compliance and verified emissions for each installation and at the country level. Emissions and compliance information on all member states’ installations under the scheme is contained in the Community Independent Transaction Log (CITL), which is publicly available and updated each year.

Emission allowances are allocated to installations for a number of phases, known as trading periods. The scheme has three trading periods, defined by their different compliance requirements. The first trading period (Phase 1) ran from 2005-2007. In order to provide flexibility for installations, the scheme
allows for intertemporal borrowing and banking of permits within each phase. That is, firms can borrow on next year’s permits to fill a shortage in the current year or bank permits for upcoming years. Inter-phase borrowing and banking was not allowable between Phases 1 and 2. However, allowances distributed during Phase 2 can be banked for use in Phase 3. Banking/borrowing provisions may have an impact on the carbon price. Banking may help installations hoping to hedge risks against seasonal and cyclical price swings. On the other hand, interperiod restrictions may severely affect the price of carbon. Alberola and Chevallier (2009b), for example, demonstrate how the restriction in banking and borrowing between Phases 1 and 2 was responsible for dramatic price falls observed in 2007.

One notable feature of Phase 2 was the introduction of carbon offsets. As an alternative to obtaining EU ETS carbon permits, a firm may offset some of its carbon emissions by investing in emission reductions elsewhere in the world. Under this directive, firms can purchase a limited number of carbon offsets called CERs (certified emission reductions) and ERUs (emission reduction units) in order to meet their obligations.

2.2 Empirical Studies of Price Drivers in the EU ETS Market

An important focus of existing studies is the role of energy markets and factors which impact on the demand or supply of energy (e.g. weather) in influencing carbon prices. Since the power sector has received more than 50% of EUAs, this focus is understandable. The energy switching behaviour of power generators is of particular interest. Power installations in the EU generate electric power largely using natural gas and coal. Operators will tend to consider the difference in relative profits from electricity generation using coal and gas. Simply by switching to gas, they can achieve some abatement of their CO₂ emissions, and as a consequence, will require fewer permits. Economic theory states that in equilibrium, the price of a permit should equal the marginal cost of abatement. In the short-term, abatement will be based on the price of switching from coal to gas. Thus the marginal abatement cost should depend on the difference between the price of gas and the price of coal. If gas prices rise, the marginal abatement cost will also rise and thus the price of a permit should as well. However, if coal prices rise such that coal generation is only a bit cheaper than gas generation, marginal abatement costs will fall, as will the price of a permit. In short, the price of coal and gas are important factors in driving carbon prices, through the ability of producers to engage in fuel switching behaviour (Christiansen et al 2005; Chevallier 2009).

Empirical studies have indeed found such associations. Keppler and Mansanet-Bataller (2009) demonstrate that during Phase 1, coal and gas prices, through clean dark and clean spark spreads⁴, impacted on CO₂ future prices. Alberola

---

⁴These spreads are the margin a plant makes from selling a unit of electricity, having bought the required fuel and carbon permits. The dark spread refers to coal and the spark
et al. (2008b) and Hintermann (2010) obtain similar findings for the spot price. Bunn and Fezzi (2007) find the carbon price reacts quickly to the gas price. Kannen (2006), Convery and Redmond (2007) and Mansanet-Batallet et al. (2010) also find that energy price changes are significantly related to carbon prices, with increases in the price of oil (which prices of natural gas are closely tied to) having the most substantial effect.

Energy demand is also influenced by weather conditions. Extremes of temperature in either direction (i.e., cold winters or hot summers) lead to higher demand for energy and, thus, affect the price of a carbon permit. Mansanet-Batallet et al. (2010) use as explanatory variables EU wide fuel prices and a weather index of several cities in the first year of the carbon market and find that the temperature in Germany is the only significant driver of carbon permit prices. Hintermann (2010) finds that temperatures affect the price of carbon more significantly in the 2006 post-crash period in permit prices. Nordic reservoir levels (which impact on the cost of generating hydroelectric power) and precipitation in Nordic countries (a proxy for reservoir levels in these areas) are also found to have negative and significant impact on carbon prices. However, their effect is diminishing, which reflects the fact that reservoir capacity is limited as is the amount of rainfall captured. Alberola et al. (2008b) focus their weather analysis on extreme weather events, i.e., hot summers and cold winters in 4 countries (Great Britain, France, Germany, Spain), weighted by the share of each country’s allocated permits. The authors find that unanticipated extremely cold (but not hot) weather events are significantly and positively related to changes in carbon prices.

Other market fundamentals, such as macroeconomic and industrial measures, have been found to be empirically associated with carbon prices. Chevalier (2009) examines the role of commonly examined macroeconomic drivers in driving futures price volatility (i.e., equity dividend yields, the junk bond premium, the US Treasury bills, and the excess return on a global commodity index). Bredin and Muckley (2010) measure the impact on carbon futures of equity prices and an index of industrial production. Both of these variables are significant in the two phases, although industrial production was found to have a counter-intuitively negative sign.

Examining the role of sectoral output in the combustion, paper and iron sectors in large allowance-holding countries (France, Germany, Italy, Poland the UK), Alberola et al. (2009a) find that output in the combustion sector in countries (except in France and Italy) significantly impacted on EUA price changes in Phase 1. But, the coefficient on output in the iron sector had a negative sign, which could be explained by the net long position of this sector (i.e., too many permits were allocated to this sector). A variable controlling for a structural break from 25 April to 23 June 2006, when prices plunged on the announcement that many countries were over allocated permits, was significant for all countries with the exception of Germany. The authors attribute the latter spread to gas. The key point to note is that they heavily reflect the price of coal or gas, respectively.
finding to the net long position of this country’s sector in this period. Alberola et al (2008a) also find that only production in combustion and iron sectors in the EU 24 were significant explanatory variables in regressions involving EUA price changes in Phase 1. More detail on the EU-ETS can be found in the online appendix.

In addition to the above market fundamentals, institutional features and events have been modelled for their potential role in affecting the price of carbon. Typically, these features are measured by the use of dummy variables. We will discuss such dummy variables in the next section.

3 Data

3.1 General Discussion of Properties of the Data

Our dependent variables are based on the spot and future prices of a carbon permit. The future has settlement date at the end of 2012 (i.e. this is the futures contract with settlement at the most distant period available in our data set). Our daily data set runs from 22 April, 2005 through 18 August, 2010.

Figure 1 plots the spot and future prices of a carbon permit. Two important aspects of the spot price series can immediately be seen. The first is evident in the middle of 2006, when the spot price dropped dramatically from roughly €30 to €10. This occurred in response to the first release of emissions accounting data for the EU ETS. This emissions verification showed that too many permits had been released in the sense that actual emissions in 2005 were well below the number of allocated permits. Subsequent to the price collapse in April 2006, a brief stabilization of the price took place until late 2006. But by early 2007 the spot price had fallen again to less than €1, with a carbon permit becoming virtually worthless (10 euro cents or less) by mid 2007.

The second feature is the sudden jump in the spot price, which occurs at the beginning of 2008, when Phase 1 of the EU ETS ended and Phase 2 began. This rapid jump in price occurred with the switch from worthless Phase 1 permits to Phase 2 permits. Recall that EU ETS carbon permits from Phase 1 could not be banked for use in Phase 2. Thus, formally speaking, in the case of the spot price the product being bought is a Phase 1 carbon permit for use in 2005-2007 and in Phase 2, a carbon permit for use in 2008-2012. For our futures price series the product being bought is always a Phase 2 carbon permit.

In order to address some of these data features, various statistical procedures are used in the literature. Some papers carry out statistical tests for structural breaks (e.g. Chevalier 2009), use estimation methods designed to be robust to structural breaks, work with sub-samples of the data, or use other methods. Others use knowledge of institutional detail of the sort described in Section 2.1 to create dummy variables. For instance, Hintermann (2010) uses an “emissions verification dummy” which equals one on 25-28 April of each year (which is the time early information was released in 2006 about the previous year’s emissions). Alberola et al (2009b) and Chevalier (2009) use a dummy variable for the entire
period, 25 April through 26 July, 2006. Alberola et al (2009b) additionally use a dummy variable for 30 March through 30 April of each year (a period which includes the yearly verification announcements). Alberola et al (2008b) include dummy variables for April 2006 and also for the period after October 2006 when a second slump in the carbon price occurs. Frunza et al (2010) argues that a period of carousel fraud had a substantial impact on carbon prices and specifies the period 1 November, 2008 through 31 August, 2009 as the time period the fraudsters were active, using a dummy variable for this period. Similarly, Mansanet-Bataller et al (2010) use a dummy variable for the credit crisis (covering the period 17 August, 2007 through 31 March, 2009) and four dummy variables that relate to news that could impact on the carbon market concerning the release of information about the allocation of permits in Phase 2, developments relating to Phase 3, and the meetings of various executive boards and other groups for making plans relating to the CER market and the linking of emission trading schemes worldwide. In short, a myriad of different dummy variables have been used by different authors.

These considerations raise the issue of whether the EU ETS provides a suitable market for forecasting or addressing important questions of interest (e.g. what are the price drivers for carbon prices?). Strong evidence exists of repeated structural change in the EU ETS carbon market due to its various institutional features. For instance, it is possible that the carousel fraud problem in May/June 2009 or the price collapse in late 2006 that carried through to 2007
caused by the overallocation of permits make these periods so different from each other (and from other times) that it is meaningless to attempt a statistical analysis which includes them all. That is, it is possible that these factors completely explain the time series dynamics in these periods and that there is no role for the price drivers to play. It also may be the case that Phase 1 and Phase 2 are so different from one another that it is meaningless to pool them both together into a single statistical analysis. But we would expect this not to be the case. Even though institutional details and other unique aspects of the carbon market undoubtedly play an important role in carbon pricing, it is likely that the price drivers discussed above as related to persistent market fundamentals, also play a role, even though their effects may be weakened or masked by institutional features and events. The challenge, taken up in this paper, is to devise a statistical methodology that can deal with this and the other issues raised above.

The existing literature exhibits a belief that it is possible to answer important empirical questions with carbon market data and attempts to address the issues noted above in various ways. One strategy in the literature is to work with futures instead of spot price data. From Figure 1, it can be seen that the futures price exhibits fewer of the problems that occur with the spot price. For instance, Bredin and Muckley (2010) use futures data and document the emergence of a stable market where prices are driven by fundamentals. However, the price of futures is not exactly the same as the spot price. The price of a future reflects investors’ beliefs about the price of a carbon permit at the settlement date. The further into the future the settlement date is, the more futures and spot price dynamics might differ.

As discussed, another strategy is to use dummy variables to control for institutional features. Since these dummies are often selected in retrospect, they cannot be used in a forecasting exercise. Furthermore, inclusion of dummy variables allows for the intercept of a regression to shift, but not the marginal effects of the price drivers. In our case, it is plausible that the marginal effects will also change in response to the factors noted above.

Another characteristic of this carbon market is that firms were able to use carbon offsets (instead of carbon permits) to partially cover their CO$_2$ emissions in Phase 2, but not in Phase 1. This raises the question of whether the set of price drivers for an EU ETS carbon permit may be different in Phase 2 than Phase 1. That is, the price drivers in Phase 1 might reflect European considerations, whereas Phase 2 price drivers could be those affecting both the international CER market and the EU ETS market. For instance, unusually hot summers which increase electricity demand would increase the demand for carbon permits in Phase 1 and Phase 2. But in Phase 2 this driver might have less of an effect on the EU ETS price since firms could choose to buy carbon offsets on the CER market, which (since it is a world-wide market) would be less affected by an unusually hot European summer. This institutional change provides a further justification for a statistical methodology that allows for the marginal effects of price drivers on the carbon price to change over time.

Another approach used by some papers for surmounting such data problems
is to use only data for Phase 1 or Phase 2 (although papers such as Bredin and Muckley 2010 do use data that span both phases). However, the restriction of the analysis to one phase omits potentially useful information in both phases about financial and other features of interest, such as the role of the price drivers. Hence a more accurate estimation of such features of interest will be obtained from using all the data. However, it is important to use a statistical methodology that allows for parameters and models to change across phases.

3.2 Data Description

Before turning to our statistical methodology, we specify and justify our set of price drivers. These are chosen as covering a wide range of issues and have all been used before in the literature (see Section 2.2). For future reference, it is worth noting that this list of predictors is quite a long one, raising statistical concerns about over-fitting.

a) Dependent variables:

The daily spot price of carbon permits (EUA) is taken from the Bluenext and Powernext exchanges. Data from the Bluenext Exchange, one of the most liquid spot market, was unavailable for the early months of 2005; hence we supplement the daily data for this period with data from the Powernext spot exchange. Measured in euros/tonne of CO$_2$.

Daily futures price of carbon permit (EUA) for settlement on December 2012, the furthest date ahead for which we had complete data for both phases. Measured in euros/tonne of CO$_2$.

b) Explanatory variables:

We use the following independent variables reflecting a wide range of macro-economic/financial and institutional variables (acronyms in brackets):

- Oil price ($p_{oil}$). Daily futures (month) ahead price of oil (brent) in euros/barrel from the Intercontinental Commodities Exchange (ICE).

- Coal price ($p_{coal}$). This is the McCloskey NW Europe Steam Coal marker.

- Gas price ($p_{gas}$). Measured as the Netherlands TTF day ahead continental gas futures negotiated in one of Europe’s largest exchanges, APX-ENDEX. Measured in euros/mwh.

- Electricity price ($p_{el}$). Measured as the Phelix base load daily one month ahead price of electricity negotiated on the European Energy exchange in euros/mwh.

- Temperature ($temp_{dev}$). The daily temperature of each EU country was calculated and a weighted average across countries was taken (weights were proportional to population). The absolute value of the deviation from this mean temperature was calculated (see Bredin and Muckley 2010). Data were from the European Climate Assessment Dataset (Klein Tank et al 2010).
Availability of hydropower energy variables (resprop) and (precip). The first of these variables was constructed from data from the Noordpool exchange which consists of the main hydropower producers in Europe (Norway, Sweden, Denmark and Finland). It was constructed by taking weekly reservoir capacity in the three countries and dividing by the total maximum possible. It was obtained from the Noordpool exchange. Reservoir capacity will affect the supply of hydroelectric power as will precipitation (Hintermann 2010). Consequently, the second variable is the precipitation level in two Noordpool countries for which complete daily data were available, Sweden and Finland.

Stock price (pstock). Daily index of the stock prices of the most highly capitalized, 100 blue-chip companies in Europe (the Euronext 100). This variable is a proxy for performance of the financial markets in the EU. It was obtained from Thompson Datastream.

Commodity price (pcomm). Daily index of world commodity prices (including energy). It was obtained from Thompson Datastream.

Corporate risk premium variable (junkprem). The daily junk bond premium is calculated as the difference between Moody’s BAA and AAA-rated bonds. The higher the spread, the greater the credit risk.

Carbon prices in the US (pcarbus). Daily carbon permit prices on the Chicago Climate Exchange (CCX).\textsuperscript{4} The variable is measured in US dollars per metric tonne of carbon.

Carbon offset price (CERfuture). We use the daily price of CER futures since 2008. We use the futures data rather than spot data on CERs since the latter has incomplete coverage. It is derived from the European Climate Exchange (ECX) and is measured as the daily CER futures price in euros with settlement in December 2012.

Overallocation of permits measure (overalloc). This variable is constructed by subtracting allocated permits from actual emissions each year.

4 Statistical Methods

The considerations of the previous sections mean that we want a statistical methodology that:

\textsuperscript{4}This is a voluntary carbon market in the US. The commodity traded is called a CFI (carbon finance instrument). Participants voluntarily agree to a legally binding commitment to cut emissions. The CCX is the largest carbon exchange in the US, and thus has the potential to affect the size of the market, if a comprehensive mandatory cap-and-trade legislation were ever to be passed. In the absence of mandatory legislation for now, a strong voluntary market in the US (strengthened by anticipation among traders of new mandatory legislation) may send a bullish signal to compliance markets elsewhere.
1. Is an extension of a regression model where a dependent variable (e.g., the price of a carbon permit) depends on a large set of explanatory variables and voids problems with over-fitting that can occur when the number of explanatory variables is large.

2. Allows for the forecasting model to change over time.

3. Allows for the marginal effects of the regression model to change over time.

4. Allows for changes in volatility (which is an important consideration for any analysis involving daily financial data).

In this paper, we use a statistical methodology called DMA, developed in Raftery et al (2010), which satisfies these characteristics and the reader is referred to Raftery et al (2010) for additional motivation and properties of DMA.

To explain DMA, we begin with the time-varying parameter (TVP) regression model:

\[
y_t = Z_{t-1} \theta_t + \varepsilon_t \\
\theta_{t+1} = \theta_t + \eta_t,
\]

where \(y_t\) is the dependent variable (e.g., the price of a carbon permit), \(Z_{t-1}\) is a \(1 \times m\) vector of observations on explanatory variables that are used for forecasting \(y_t\), \(\theta_t\) is an \(m \times 1\) vector of regression coefficients, \(\varepsilon_t\) is \(N(0, H_t)\) and \(\eta_t\) is \(N(0, Q_t)\). Note that we are forecasting one day ahead and, thus, \(Z_{t-1}\) is the information available for forecasting \(y_t\) (and in our empirical work, our price drivers will always be lagged one day). The TVP regression model is a state space model of the sort commonly used in empirical macroeconomics (see, e.g., among many others, Cogley and Sargent 2005; Cogley, Morozov and Sargent 2005; Primiceri 2005). Standard methods (e.g., involving the Kalman filter) for estimation and prediction exist with such models. They allow for the marginal effects of the predictors to change over time. This is potentially very useful with the present data set. For instance, if the gradual decline of the carbon permit price in late 2006 was associated with a gradual change in the role of some of the price drivers, then a TVP regression model would be a suitable way of modelling this.

However, TVP regression models can deal poorly with the sort of abrupt changes in the time series dynamics that we may expect with this data set. Furthermore, they can be over-parameterized. That is, if \(m\) is moderately large (as

\[5\]Note that some of the carbon market literature uses cointegration methods. However, this is unnecessary here. When working with TVP models, researchers typically do not worry about unit root or cointegration issues. TVP models are very flexible and their focus is on fitting the time variation in coefficients. Even if variables have unit roots this is not a problem in the sense that all that matters is that \(\varepsilon_t\) and \(u_t\) do not have unit roots. The time-varying \(\theta_t\) can effectively pick out cointegrating relationships (if they exist) leading to stationary residuals. Even if they fail to do so, note that the presence of a time-varying intercept which follows a random walk will pick up any remaining nonstationarity in the model.
in the present application) and the data are noisy, even a standard regression model can over-fit. In regression problems with many potential explanatory variables, Bayesian model averaging (BMA) has been a common response to such worries (see, among many others, Raftery, Madigan and Hoeting, 1997, Fernandez, Ley and Steel 2001a,b; Eicher, Papageorgiou and Raftery 2010 and Ley and Steel 2010). However, until very recently, it has been considered computationally infeasible to do BMA with TVP regressions such as (1). The contribution of Raftery et al (2010) was to develop computationally feasible methods of doing model averaging with dynamic models such as (1). The empirical work in Raftery et al (2010) involves an engineering application, but Koop and Korobilis (2011) find DMA to work well in a macroeconomic application.

To define the set of models used in DMA, let \( Z_t^{(k)} \) contain a subset of the price drivers in \( Z_t \) for \( k = 1, \ldots, K \). Each of the models in our model space is:

\[
\begin{align*}
    y_t &= Z_{t-1}^{(k)} \theta_t^{(k)} + \varepsilon_t^{(k)} \\
    \theta_{t+1}^{(k)} &= \theta_t^{(k)} + \eta_t^{(k)},
\end{align*}
\]

\( \varepsilon_t^{(k)} \) is \( N \left( 0, H_t^{(k)} \right) \) and \( \eta_t^{(k)} \) is \( N \left( 0, Q_t^{(k)} \right) \). Let \( L_t \in \{1, 2, \ldots, K\} \) denote which model applies at \( t \) and \( y_t = (y_{1t}, y_{2t}, \ldots, y_{mt})' \). For notational simplicity, we do not explicitly include \( Z_{t-1} \) in the conditioning argument in the densities below (but all densities are conditional on the price drivers). The fact that we are letting different models hold at each point in time and will carry out model averaging justifies the terminology “dynamic model averaging”. To be precise, DMA involves estimating \( \theta_t^{(k)} \) in each individual model and averaging these model-specific estimates across the model space using \( \Pr \left( L_t = k | y_{t-1} \right) \) for \( k = 1, \ldots, K \) as weights. When forecasting, DMA involves averaging across predictive densities or point forecasts using \( \Pr \left( L_t = k | y_{t-1} \right) \) for \( k = 1, \ldots, K \) as weights. Dynamic model selection (DMS) can be done by forecasting using only model \( j \) where \( \Pr \left( L_t = j | y_{t-1} \right) \) has the maximum model probability.

Since \( Z_t \) contains \( m \) price drivers and there are \( 2^m \) possible subsets of these, we have \( K = 2^m \) models. Note that, in a recursive forecasting exercise with such a set of TVP models, a computational challenge arises since a different forecasting model may apply at each point of time and the number of combinations of models which must be estimated in order to forecast at time \( t \) is \( 2^{mr} \). For empirically-relevant choices of \( m \) and \( r \) it is computationally infeasible to use MCMC methods to evaluate all these TVP models. The contribution of Raftery et al (2010) is to use forgetting factors (described below) to develop computationally efficient approximations so as to allow us to do DMA.

The advantage of DMA is that it allows for switches between parsimonious models. That is, instead of suffering the over-fitting problems that often occur with TVP regression models that include all \( m \) price drivers at each point in time (or suffering from misspecification problems that arise if constant coefficients models or models with too few price drivers are used), DMA can switch from (say) a TVP model with three or four price drivers to a TVP model with three
or four different price drivers. In practice (see, e.g., Koop and Korobilis 2011) DMA has been found to favour parsimonious models and can handle abrupt changes in the dynamic structure (such as might have happened when Phase 1 ended and Phase 2 begins) much better than conventional TVP models.

We work with the set of models defined in (2). Estimation within a single model involves textbook statistical methods involving the Kalman filter and will not be explained here in detail. Apart from our estimation of \( H_t^{(k)} \), which will be discussed below, our estimation methods within a single model are exactly as in Raftery et al (2010).

To briefly define summarize the DMA treatment of a single model, note that Kalman filtering within model \( k \) is based on two main equations:

\[
\theta_{t-1}^{(k)} | y^{t-1} \sim N \left( \tilde{\theta}_{t-1}^{(k)}, \Sigma_{t-1|t-1}^{(k)} \right)
\]

and

\[
\theta_t^{(k)} | y^{t-1} \sim N \left( \tilde{\theta}_t^{(k)}, \Sigma_{t|t-1}^{(k)} \right),
\]

where the formulae for \( \tilde{\theta}_t^{(k)} \) and \( \Sigma_{t|t-1}^{(k)} \) have standard textbook forms and

\[
\Sigma_{t|t-1}^{(k)} = \Sigma_{t-1|t-1}^{(k)} + Q_t^{(k)}.
\]

Beginning at \( t = 0 \), Kalman filtering sequentially updates these formulae and forecasting is done using the predictive distribution

\[
y_t | y^{t-1} \sim N \left( Z_{t-1} \tilde{\theta}_{t-1}^{(k)}, H_t^{(k)} + Z_{t-1} \Sigma_{t|t-1}^{(k)} Z_t^{(k)^T} \right).
\]  

(3)

Note that these formulae depend on \( Q_t^{(k)} \) and \( H_t^{(k)} \) and conventional Bayesian estimation of them (e.g. using stochastic volatility or GARCH specifications) will typically require the use of MCMC methods which is computationally infeasible in the context of DMA. Accordingly, Raftery et al (2010) use a forgetting factor approach where:

\[
\Sigma_{t|t-1}^{(k)} = \frac{1}{\lambda} \Sigma_{t-1|t-1}^{(k)}
\]

or, equivalently, \( Q_t^{(k)} = (1 - \lambda^{-1}) \Sigma_{t-1|t-1}^{(k)} \) where \( 0 < \lambda \leq 1 \). Such approaches have long been used in the state space literature. The name “forgetting factor” is suggested by the fact that this specification implies that observations \( j \) periods in the past have weight \( \lambda^j \). Following Raftery et al (2010), we set \( \lambda = 0.99 \).

With regards to initialization of the Kalman filter, we use a diffuse prior \( \theta_0^{(k)} \sim N \left( 0, 100I_{m_k} \right) \), where \( m_k \) is the number of variables in model \( k \), for \( k = 1, \ldots, K \).

The contribution of Raftery et al (2010) was to derive a forgetting factor approach to carry out DMA. To simplify notation, let \( \pi_{t|s,k} = \Pr \left( L_t = k | y^s \right) \). The new recursions required by DMA involve (beginning with \( \pi_{0|0,k} \)):

\[
\pi_{t|t-1,k} = N \left( \tilde{\pi}_{t|t-1,k}, \Sigma_{t|t-1,k} \right)
\]
and \( \pi_{t|t,k} \). DMA proceeds by averaging across forecasts using \( \pi_{t|t-1,k} \) as weights for \( k = 1, \ldots, K \) and \( t = 1, \ldots, T \). For instance, DMA point forecasts are given by:

\[
E (y_t|y^{t-1}) = \sum_{k=1}^{K} \pi_{t|t-1,k} Z^{(k)}_{t-1} \tilde{\theta}^{(k)}_{t-1},
\]

where \( \tilde{\theta}^{(k)}_{t-1} \) is the Kalman filter estimate of the regression coefficients given information available at time \( t-1 \).

Raftery et al (2010) use:

\[
\pi_{t|t-1,k} = \frac{\pi_{t-1|t-1,k}^{\alpha}}{\sum_{l=1}^{K} \pi_{t-1|t-1,l}^{\alpha}},
\]

where \( 0 < \alpha \leq 1 \) is a forgetting factor which is set to a fixed value slightly less than one. The huge advantage of using the forgetting factor in the model prediction equation is that we do not require an MCMC algorithm to draw transitions between models nor a simulation algorithm over model space. Instead, simple evaluations comparable to those of the Kalman filter can be done. In particular, we have a model updating equation of:

\[
\pi_{t|t,k} = \frac{\pi_{t-1|t-1,k} P_k \left( y_t|y^{t-1} \right)}{\sum_{l=1}^{K} \pi_{t-1|t-1,l} P_l \left( y_t|y^{t-1} \right)},
\]

where \( P_k \left( y_t|y^{t-1} \right) \) is the predictive density for model \( k \) given in (3) evaluated at \( y_t \).

To understand further how the forgetting factor \( \alpha \) can be interpreted, note that this specification implies that the weight used in DMA which is attached to model \( k \) at time \( t \) is:

\[
\pi_{t|t,k} \propto \left[ \pi_{t-1|t-2,k} P_k \left( y_{t-1}|y^{t-2} \right) \right]^{\alpha} \prod_{i=1}^{t-1} \left[ P_k \left( y_{t-i}|y^{t-i-1} \right) \right]^{\alpha^i}.
\]

Thus, model \( k \) will receive more weight at time \( t \) if it has forecast well in the recent past (where forecast performance is measured by the predictive density, \( P_k \left( y_{t-i}|y^{t-i-1} \right) \)). The interpretation of “recent past” is controlled by the forgetting factor, \( \alpha \) and we have an exponential decay at the rate \( \alpha^i \) for observations \( i \) periods ago. Thus, if \( \alpha = 0.99 \) (the value used by Raftery et al, 2010 and in this paper), forecast performance four weeks ago receives 80% as much weight as forecast performance last period (when using daily data based on a five day week).

We stress that, conditional on \( H_t \), the estimation and forecasting strategy outlined above only involves evaluating formulae such as those in the Kalman filter. The recursions above are started by choosing \( \pi_{0|0,k} \) for \( k = 1, \ldots, K \) and we make the noninformative choice of \( \pi_{0|0,k} = \frac{1}{K} \) for \( k = 1, \ldots, K \).
The preceding discussion is all conditional on $H_t$. Raftery et al (2010) recommend a simple plug in method where $H_t^{(k)} = H^{(k)}$ and is replaced with a consistent estimate. When forecasting financial variables, however, it is likely that the error variance is changing over time. Thus, we use an Exponentially Weighted Moving Average (EWMA) forecast of $H_t^{(k)}$:

$$\hat{H}_t^{(k)} = \kappa \hat{H}_{t-1}^{(k)} + (1 - \kappa) \left( y_{t-1} - Z_{t-2} \beta_{t-1} \right)^2.$$ 

EWMA estimators are commonly used to model time-varying volatilities in finance; see Riskmetrics (1996) for the properties of EWMA estimators. $\kappa$ is called a decay factor, and Riskmetrics proposes setting 0.94 for daily data and we follow this choice.

5 Forecasting Carbon Prices in the EU ETS

5.1 Introduction

All the explanatory variables in our models are lagged by one day relative to the dependent variable. All of the models include an intercept and a lag of the dependent variable. Our models differ in their treatment of the 13 price drivers given in Section 3.2. Thus, at each point in time, DMA is model averaging over $K = 2^{13} = 8192$ models.

We divide our results in two parts: one applies to estimation using DMA; the other to forecasting and includes a comparison with a variety of different approaches. All variables are logged except for those which take on zero or negative values. We repeat all our empirical work four times: using two dependent variables (i.e. the logs of the spot and future prices) and two samples (i.e. the entire sample and only Phase 2). For our recursive forecasting exercise, all of these four cases our discussed. For our discussion of price drivers, which involve estimation using the entire sample, we present a selection of representative results, focussing mainly on futures data. Complete results are available in the online appendix.

5.2 Price Drivers

To summarize some main features of our data, Table 1 presents OLS estimates the regression using all the price drivers. It can be seen that, apart from the lag of the dependent variable, there is no price driver which is consistently significant at the 5% level in all data sets and time periods. However, there are many variables which are significant in some cases and many t-statistics between 1 and 2 that would be significant at lower significance levels. This is what one would expect to find when one has many explanatory variables which are highly correlated with one another and their coefficients are changing over time, possibly being significant using some sub-sets of the data, but not in
others. Such considerations help motivate our use of DMA and, hence, we turn to our results using DMA.

Table 1: OLS Results Using Various Data Sets (t-stats in parentheses)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Spot Phase 1+2</th>
<th>Future Phase 1+2</th>
<th>Spot Phase 2</th>
<th>Future Phase 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.628 (0.989)</td>
<td>0.088 (1.134)</td>
<td>-3.920 (-12.017)</td>
<td>-0.029 (-0.247)</td>
</tr>
<tr>
<td>Lag of dep var</td>
<td>0.989 (138.209)</td>
<td>0.977 (149.344)</td>
<td>-0.220 (-9.672)</td>
<td>0.977 (72.868)</td>
</tr>
<tr>
<td>poil</td>
<td>0.226 (1.473)</td>
<td>0.005 (0.226)</td>
<td>-0.034 (-0.451)</td>
<td>-0.023 (-0.825)</td>
</tr>
<tr>
<td>pcoal</td>
<td>-0.002 (-0.032)</td>
<td>0.009 (1.547)</td>
<td>-0.139 (-3.55)</td>
<td>0.034 (2.544)</td>
</tr>
<tr>
<td>pgas</td>
<td>0.057 (1.586)</td>
<td>-0.006 (-1.352)</td>
<td>-0.015 (-0.560)</td>
<td>-0.029 (-3.029)</td>
</tr>
<tr>
<td>pel</td>
<td>-0.010 (-0.220)</td>
<td>0.006 (1.361)</td>
<td>0.061 (2.574)</td>
<td>0.002 (0.968)</td>
</tr>
<tr>
<td>pstock</td>
<td>-0.158 (-1.955)</td>
<td>-0.016 (-1.604)</td>
<td>0.7281 (16.027)</td>
<td>0.004 (0.270)</td>
</tr>
<tr>
<td>junkprem</td>
<td>0.017 (0.698)</td>
<td>-0.004 (-1.520)</td>
<td>0.0153 (1.558)</td>
<td>-0.002 (-0.570)</td>
</tr>
<tr>
<td>resprop</td>
<td>-0.052 (-1.537)</td>
<td>-0.007 (-1.595)</td>
<td>0.111 (5.344)</td>
<td>-0.002 (-0.273)</td>
</tr>
<tr>
<td>tempdev</td>
<td>0.007 (1.246)</td>
<td>-4.5×10^{-4} (-0.561)</td>
<td>-0.033 (-10.770)</td>
<td>-0.002 (-2.164)</td>
</tr>
<tr>
<td>precip</td>
<td>-2.3×10^{-6} (0.112)</td>
<td>-2.3×10^{-6} (0.950)</td>
<td>5.3×10^{-9} (0.962)</td>
<td>-2.3×10^{-9} (-0.119)</td>
</tr>
<tr>
<td>pcomm</td>
<td>-0.121 (-0.761)</td>
<td>0.006 (0.288)</td>
<td>0.274 (3.482)</td>
<td>0.019 (0.703)</td>
</tr>
<tr>
<td>pcabus</td>
<td>-0.010 (-1.021)</td>
<td>0.002 (2.156)</td>
<td>0.051 (11.636)</td>
<td>0.003 (1.741)</td>
</tr>
<tr>
<td>overalloc</td>
<td>-4.8×10^{-9} (-0.386)</td>
<td>-4.6×10^{-9} (-0.271)</td>
<td>3.5×10^{-9} (6.370)</td>
<td>1.8×10^{-8} (0.962)</td>
</tr>
<tr>
<td>cerfuture</td>
<td>-0.032 (-1.654)</td>
<td>-0.004 (-1.878)</td>
<td>0.687 (13.947)</td>
<td>-0.010 (-1.175)</td>
</tr>
</tbody>
</table>

Figures 2 through 5 summarize the information provided by our many models and parameters when the log of the carbon future is the dependent variable and we use the entire data set. Figure 2 sheds light on how parsimonious DMA is. Let $Size(k)$ be the number of price drivers in model $k$ then

$$E(\text{Size}_t) = \sum_{k=1}^{K} \text{Pr}(L_t = k|y^{t-1}) \times \text{Size}_t$$

can be interpreted as the expected or average number of price drivers used by DMA at time $t$. Figure 2 indicates that DMA is wanting to use roughly half of
the 13 price drivers for most of the time span of our data and, thus, is achieving a fair degree of parsimony relative to a TVP regression model.

Figures 3 through 5 present what we call the “posterior importance” of each variable. The posterior importance of the $j^{th}$ price driver is defined as the probability DMA attaches to models that include the $j^{th}$ price driver.$^6$

In general, Figures 3 through 5 show that constant coefficient models that simply include all the price drivers are inappropriate. Figures 3 through 5 demonstrate that there is a great deal of variation over time in respect to which price drivers are included.

We hesitate to tell strong stories about individual price drivers in the context of a reduced-form forecasting exercise such as the present one. Nevertheless, a few patterns are worth noting. It is rare for DMA to attach a probability close to one to any particular price driver at any point in time. This is to be expected with a financial forecasting exercise (where often the first lag of the dependent variable provides most of the predictive power) with correlated explanatory variables. As with previous studies, though, the prices of gas, oil, coal and electricity are at some points in time important price drivers. Note, in particular, the increasing role of natural gas prices in 2009.

$^6$For the sake of brevity, we do not present a graph for the CERfuture variable since it is found to play relatively little role.
In terms of financial variables, the junk bond premium (a measure of financial risk) is sometimes important and briefly becomes very important at the height of the financial crisis in the autumn of 2008. Figure 5 shows that the US carbon market variable (pcarbus) is a moderately strong price driver over the whole data span.

![Figure 3: Results for Carbon Futures Data](image-url)
Figure 4: Results for Carbon Futures Data
Figures 3 through 5 indicate which price drivers are important (and when), but they do not tell the magnitude of the effects of each price driver on the carbon future price. With so many price drivers, all having time varying coefficients, presenting coefficient estimates and credible intervals would take up a great deal of space. The interested reader is referred to the online appendix which presents this information for all of the coefficients. It also provides a complete set of results for the other data sets and sub-samples. The online appendix also plots the error variances. Suffice to note that there is strong evidence for time variation in the error variance. In particular, it shows a large increase in two time periods. These correspond to the spring 2006 when the initial emissions verification process revealed the over-allocation of permits; and early 2009, when fraudulent activity affected the market (Frunza et al 2010).

5.3 Results of Forecasting Exercise

To evaluate the forecasting performance of DMA in the carbon market, we use mean squared forecast errors (MSFEs) and sums of log predictive likelihoods. The MSFE is a standard measure of the performance of point forecasts, whereas the sum of log predictive likelihoods evaluates the forecast performance of the entire predictive density. Formally, the predictive likelihood at time $t$ is the predictive density (given information through $t - 1$) evaluated at the actual outcome. DMA provides us with the predictive density as described in Section...
4. For both MSFEs and sums of log predictive likelihoods we set aside an initial 100 daily observations and evaluate forecast performance on the remaining observations.

We compare DMA and DMS to several alternatives with short-form names/acronyms: i) BMA, ii) Bayesian model selection (BMS), iii) time-varying parameter regression model (TVP), iv) constant coefficient noninformative prior (CC-NONINF) and v) constant coefficient model with shrinkage (SHRINK).

These can be considered as special cases of DMA or DMS. BMA is the same as DMA except that the time-variation in coefficients is removed. That is, in (1) it sets \( Q_t = 0 \) so that coefficients do not change over time. Otherwise, all the assumptions are the same as in DMA. Hence, this approach does model averaging using conventional constant coefficient models but the weights in the model averaging are calculated using predictive likelihoods just as is done in DMA. Similarly, BMS is the same as DMS but in a constant coefficient model. Hence, the single best model is used to forecast at each point in time, where the best model is selected using past predictive performance in the same way as in DMS.

The TVP regression model includes all the price drivers. This is one of the 8192 models used in DMA and is estimated in the same manner as when we are doing DMA. However, no model averaging is involved, but rather this single model is used.

CC-NONINF is the single regression model which includes all the price drivers and does not allow for time variation in coefficients (i.e. since we are using a noninformative prior, it is analogous to a recursive OLS forecasting strategy using a regression model containing all the explanatory variables). SHRINK is the same as CC-NONINF, however we use an informative \( N(0, I) \) prior on the regression coefficients so as to shrink them towards zero.

All of these approaches allow for the error variance, \( H_t \), to vary over time in exactly the same manner using the EWMA estimator described in Section 4.

To aid in interpretation, we present our results relative to a benchmark: the constant coefficient regression model with noninformative prior. MSFEs (sums of log predictive likelihoods) are divided by (subtracted from) those of CC-NONINF. This implies lower (more negative) values of MSFEs (log predictive likelihoods) indicate better forecast performance.

Table 2 compares the forecast performance of these different approaches using both the entire sample and using only Phase 2 data. We organize our discussion of results around several questions.

Is it important to allow for model change when forecasting with these data sets? The answer to this is clearly yes. DMA, DMS, BMA and BMS are the four approaches which allow for model change. The best-forecasting model is always chosen among this set and often their forecast performance is much better than the alternatives.

Is it important to allow for parameter change (though time variation in coefficients)? Here the evidence is weaker. BMA and BMS (which are the variants of DMA and DMS which do not allow for parameter change), typically forecast roughly as well as DMA and DMS. Clearly, it is model change which is driving
forecast improvements as opposed to parameter change.

How well would the forecaster using (increasingly popular) TVP regression models perform? The answer to this is: very poorly indeed. This approach consistently forecasts the worst and seems to be seriously over-parameterized.

Are there differences between the findings of MSFEs and predictive likelihoods? The general patterns provided by these two forecast metrics are similar to one another, but the strength of the findings is very different. Over the entire sample, MSFEs indicate that DMA forecasts only slightly better than a standard regression model (with noninformative prior and heteroskedastic errors). However, predictive likelihoods indicate much more substantial forecast improvements due to use of DMA (or any of the other approaches which allow for model change). This suggests DMA is offering modest improvements in producing point forecasts, but larger improvements in accurately modelling the shape of the entire predictive distribution.

Are there differences between spot and futures forecasts? Using the entire sample, DMA is more useful when forecasting futures data. When forecasting spot prices, it offers only small improvements relative to a constant coefficient model.

Are there differences between forecast performance over phases? The benefits of using DMA are most substantive in Phase 2. With both spot and futures data, the improvements provided by DMA (or DMS, BMA or BMS) are most striking in Phase 2.

Are there substantial differences between model averaging and model selection approaches? The answer to this is no. There are clearly large benefits to allowing for model change and, hence, the approaches which allow for this (DMA, DMS, BMA and BMS) forecast better than those which do not (TVP, CC-NONINF and SHRINK). However, model selection and model averaging approaches forecast similarly.
To understand how the results in Table 2 arise, it is worth studying the forecast performance over time in more detail. One issue is how DMA handled the abrupt break in the spot price when we switched from Phase 1 to Phase 2. When forecasting the spot price on 1 January, 2008, the real-time forecaster only has information available through 31 December, 2007. On 31 December, 2007 the log of the spot price was -3.91, before switching to 3.18, 3.20, 3.20 and 3.21 on the first four trading days of January 2008. The DMA point forecasts on these five days are -3.94, -4.11, 10.24, 3.27 and 3.20. In other words, DMA (unsurprisingly) forecasts very poorly on the first two trading days of January, but by the third day was forecasting reasonably well and by the fourth and beyond was exhibiting very good forecast performance. We find this strong evidence that DMA can effectively handle even large structural breaks in financial time series. That is, it only took DMA three or four days to adjust its forecast performance to a break of huge magnitude in the dependent variable.

Table 2 is based on sums of log predictive likelihoods over the entire sample. Figures 6 and 7 plot cumulative sums up to each point in time labelled on the X-axis for the futures and spot prices, respectively. A straight line in such a graph indicates that forecast performance is of a constant quality throughout the sample. For the futures data, apart from a small deterioration in forecast performance in May 2006 (i.e. the time when prices declined after the initial emissions verification revealed an over-allocation of permits), Figure 6 reveals such a straight line. For the spot price data, a clear deterioration in forecast performance is observed both in May 2006 and in early 2008 after the switch between Phase 1 and Phase 2.
Figure 6: Results for Carbon Futures Data
5.4 Further Discussion of Results

The DMA methodology employed in this study has revealed a number of characteristics pertinent to understanding the properties of the carbon futures and spot markets. First, the fact that so much model change is occurring suggests that this market is unstable and immature. By way of explanation, note that financial theories suggest that the price of an asset should reflect all available information about market fundamentals. Price changes should only occur in response to new information about the asset’s fundamental value. In our study, the price drivers should measure new information. We are finding that the price drivers often do have forecasting power. However, in stable, mature and efficient markets, we would expect the role of the price drivers to be roughly constant over time. For example, in mature markets we might expect that every time the coal price rises it will impact the carbon permit price in roughly the same way. We are not finding this stability in our results. The fact that statistical methodologies that allow for model change produce even more substantial improvements in forecast performance in Phase 2 than over the entire sample suggests that markets are not becoming more stable.

Furthermore, results for the spot and future prices exhibit similar levels of instability. This is not what we would expect in mature markets. That is, financial theory suggests that futures (with settlement date December 2012)
should be less sensitive to news that has a short run impact on the demand for carbon permits. However, we find, for example, that some of price drivers reflecting weather conditions (which would be expected to have most impact on the current year’s demand for carbon permits) are important for the futures price even in Phase 1. This instability in the EU ETS could have impacts on emitters, making it more difficult for them to plan ahead and achieve reductions in CO₂ emissions in a cost effective manner (e.g. through investments in low-carbon technologies and optimal design of policies programs aimed at emissions reduction). Nonetheless, we do see how a method like DMA can adapt the forecasting model quite quickly over time to handle the instabilities in the carbon market. Table 2 and Figures 6 and 7 show DMA to forecast well, with only brief deteriorations in forecast performance at the most turbulent times.

6 Conclusion

This paper uses a technique called DMA to forecast spot and futures prices in the EU ETS carbon market. Forecasting prices as accurately as possible is necessary for installations subject to compliance to make appropriate investment decisions based on price expectations and for policymakers to design appropriate emission reductions policies and calculate the real costs of emissions reduction to society. We have argued that DMA is ideally suited for studying such a market, since it deals with problems caused by the proliferation of price drivers and allows for the changing effects of price drivers in the market and for the forecasting model itself to change over time. We find strong evidence of substantial turbulence and change in the EU ETS market. We show how our DMA approach can model these features and forecast accurately compared to other approaches.
References


