

# Forecasting the European Carbon Market<sup>1</sup>

Gary Koop & Lise Tole<sup>2</sup>  
Department of Economics  
University of Strathclyde

February 2011, revised August 2011 and March 2012

**Abstract:** In an effort to meet its obligations under the Kyoto Protocol, the European Union has introduced a cap-and-trade scheme where mandated companies are allocated permits to emit CO<sub>2</sub>. Financial markets have developed that allow companies to trade these carbon permits. Several recent studies have attempted to model the price of a carbon permit. There are many institutional features that potentially impact on carbon prices (and associated financial futures), making such an undertaking quite different from modelling 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 carbon markets. We find 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

---

<sup>1</sup>We 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.

<sup>2</sup>Address 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 Emissions Trading Scheme (EU ETS), commenced operation on January 1st, 2005. The scheme caps CO<sub>2</sub> emissions and then distributes allowances to emit (carbon permits) to companies that are large energy users. These carbon permits can be traded on financial markets. Companies whose CO<sub>2</sub> emissions exceed their allowable limit can purchase additional permits in these markets. Companies with permits in excess of their emissions can sell them. 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 price of a carbon permit. But the exact list of price drivers and the magnitude of their impacts on the carbon price is unclear. Moreover, unlike other commodity markets, the EU ETS is influenced by compliance. The supply of permits is set by the European Commission (EC), which also administers the program and determines its rules and regulations. Elements of the trading scheme have changed over time. Uncertainty over a global climate agreement, the size of the future cap, and the eligibility and number of some types of offsets allowable in the scheme are also key factors impacting on the price of carbon. While all commodities markets face uncertainties, these institutional characteristics are unique to the EU ETS.

The complexities of this market raise particular challenges for the statistician. As with many financial studies (e.g. Avramov 2002), the number of potential variables that may affect carbon 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, among many others, Avramov, 2002, Sala-i-Martin, Doppelhofer and Miller, 2004, Ley and Steel, 2009 and 2010). 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 price drivers, the impacts of which will vary over time. 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. Moreover, it 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 that have appeared 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.

The remainder of this paper is organized as follows: Section 2 provides an

overview of the EU ETS scheme and reviews 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](http://personal.strath.ac.uk/gary.koop/koop_tole_appendix.pdf) provides additional discussion of institutional details, the data set, and further empirical results.

## 2 Carbon Prices in the EU ETS

### 2.1 Institutional Aspects

The EU ETS is the key institution whereby EU countries aim to meet their Kyoto obligations. 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<sub>2</sub> emissions. At the end of each year, these companies must surrender a number of carbon permits equal to the amount of their total emissions for the year. The carbon permit is called an EU allowance (EUA) and is equal to one metric tonne of CO<sub>2</sub>. EUAs can be bought and sold in financial markets. Each year the EC discloses information on compliance and verified emissions for each installation as well as emissions 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 currently has three trading periods, each defined by its specific compliance requirements. The first trading period (Phase 1) ran from 2005-2007. Phase 2 ran from 2008-2012. Phase 3 will run from 2013-2020. 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. The EC prohibited inter-phase borrowing and banking between Phases 1 and 2. However, it does permit allowances distributed during Phase 2 to be banked for future phases. 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, inter-phase 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 in the scheme. As an alternative to obtaining EUAs, a firm may offset some of its carbon emissions by investing in emission reductions elsewhere. 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

Most existing studies investigate the influence of factors that impact on the demand or supply of energy on carbon prices. Since the power sector has received more than 50% of EUAs, this focus is understandable. The fuel switching behaviour of power generators is of particular interest. Power installations in the EU generate electricity largely by using natural gas and coal. By switching to gas, they can achieve some abatement of their CO<sub>2</sub> emissions. Economic theory tells us that the price of a permit should equal the marginal cost of abatement. This marginal abatement cost will depend on the difference between the prices of gas and coal. If gas prices rise, the costs of fuel switching and thus the marginal abatement cost will also rise. Thus, the price of a permit should as well. However, if coal prices rise, marginal abatement costs will fall, as will the price of a permit. In short, the relative prices of coal and gas are important factors in driving carbon prices, operating as they do through the ability of producers to engage in fuel switching behaviour (see, e.g., Christiansen et al, 2005 and Chevallier, 2009).

Empirical studies have indeed found such associations. Keppler and Mansanet-Batallet (2009) find that during Phase 1, coal and gas prices impacted on CO<sub>2</sub> future prices. Alberola et al (2008b) and Hintermann (2010) obtain similar findings for the spot price. Bunn and Fezzi (2007) find that the carbon price reacts quickly to the gas price. Kanen (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 (to which natural gas prices are closely tied) 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) measure EU wide fuel prices and a weather index of several cities in the first year of the carbon market to analyze carbon permit prices. They find that the temperature in Germany is the only significant driver of prices. Hintermann (2010) finds that temperature significantly affects the price of carbon, particularly after April 2006. He also finds that 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) have significantly negative impacts on carbon prices. Alberola et al (2008b) focus their weather analysis on extreme weather events and find that unanticipated extremely cold (but not hot) weather events are significantly and positively related to changes in carbon prices.

Various economic and financial factors also have been found to be empirically associated with carbon prices. Chevallier (2009) included several macroeconomic risk drivers in an analysis of the macroeconomic determinants of EUA

prices, but finds only a weakly significant association between carbon prices and stock and bond variables. Bredin and Muckley (2010) find significant relationships between carbon prices and stock prices and an index of industrial production. Alberola et al (2008a, 2009a) include price drivers reflecting industrial production and allocation of permits, finding that industrial production in several high allowance sectors has a strong impact on EUA prices.

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. Additional details about the EU ETS can be found in the online appendix.

## 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 from roughly €30 to €10. This occurred in response to the first release of emissions accounting data for the EU ETS. This emissions verification data showed that too many permits had been released; that is, actual emissions in 2005 were 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. Chevallier, 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 Chevallier (2009) use a dummy variable for a longer period after the release of emissions accounting data, 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 that 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) argue that a period of carousel fraud had a substantial impact on carbon prices and use a dummy variable for 1 November, 2008 through 31 August, 2009 when the fraudsters were active. 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.

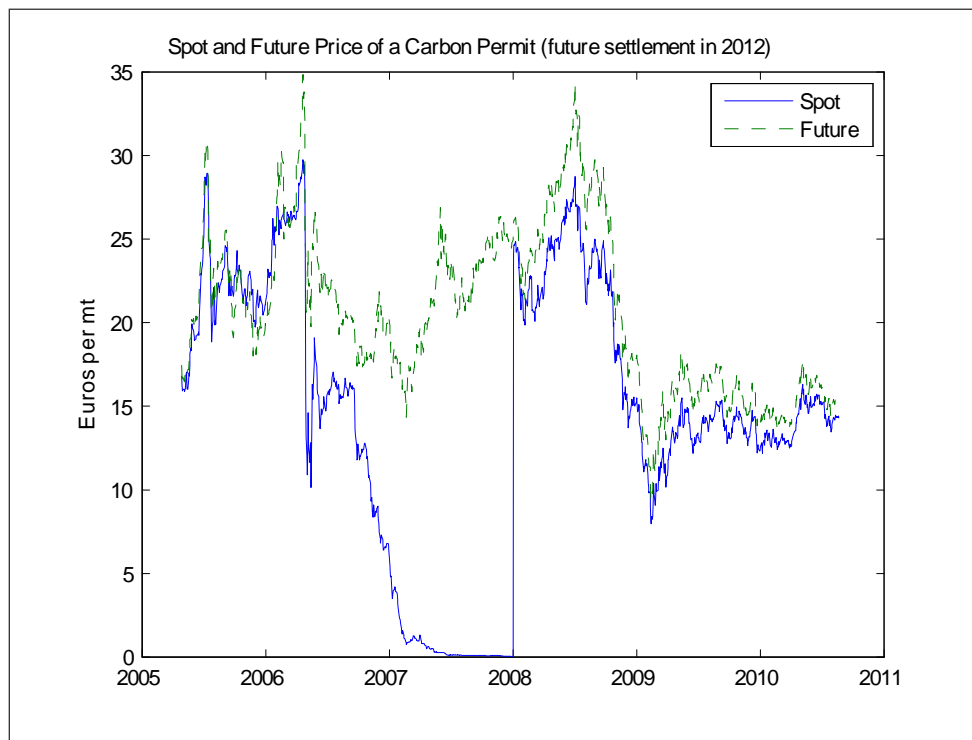


Figure 1

These considerations raise the issue of whether the EU ETS provides a suitable market for forecasting or addressing important questions of interest (i.e. 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 that includes them all. That is, it is possible that these factors completely explain the time series dynamics in these periods such 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 either 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 issue.

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, the 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<sub>2</sub> 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 European 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 EU ETS carbon permit 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 (e.g. Alberola et al. 2008b) to surmount such problems is to include 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 might 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. We follow choices made in the papers surveyed in Section 2.2. In particular, we choose variables that have been identified in the literature as potentially influencing the demand for permits (e.g. variables measuring weather conditions, energy use, and the state of the economy and investment climate that will impact on production and thus output).

a) **Dependent variables:**

The daily spot price of carbon permits (EUA) is taken from the Bluenext and Powernext exchanges. Data from the Bluenext Exchange was unavailable for the early months of 2005; hence we supplement the daily data for this period with data from the Powernext spot exchange. It is measured in euros/tonne of CO<sub>2</sub>.

The daily futures price of carbon permits (EUA) for settlement on December 2012. This is the settlement data for which the most data were available for both phases. It is measured in euros/tonne of CO<sub>2</sub>.

b) **Explanatory variables:**

We use the following independent variables.

- Oil price. Daily futures (month) ahead price of oil (Brent) in euros/barrel from the Intercontinental Commodities Exchange (ICE) where it has been traded since 2005. The Brent oil measure is the key oil marker for European oil. Hintermann (2010) notes that, while the price of coal and gas will influence the price of carbon through the fuel switching behaviour of power providers, this behaviour will not be influenced to any extent by the price of oil since the power installations in the EU depend overwhelmingly on coal and/or natural gas to produce electricity. Nevertheless, the price of oil been found to be a key price determinant of carbon prices in other studies (e.g. Mansanet-Bataller et al, 2007, Alberola et al, 2008b and Kanen, 2006). This may reflect the fact that oil prices tend to rise and fall



with other energy prices. Alternatively, a change in the price of oil may indirectly affect the price of carbon through its affect on the overall price level in the macroeconomy. Rising prices may dampen investment and raise unemployment, thereby leading to a fall in the demand for permits.

- Coal price. This is the McCloskey NW Europe Steam Coal marker. It is the major coal price indicator for Europe and should affect carbon prices through the fuel switching behaviour of electricity generating plants (e.g. Mansanet-Bataller et al, 2007)
- Gas price. Measured as the Netherlands TTF day ahead continental gas futures negotiated in one of Europe's largest exchanges, APX-ENDEX. Measured in euros/mwh. As with the price of coal, the gas price should influence the fuel-switching behaviour of the power installations (e.g. Mansanet-Bataller et al, 2007 and Kanen, 2006).
- Electricity price. Measured as the Phelix base load daily one month ahead price of electricity negotiated on the European Energy exchange in euros/mwh. Given that a large majority of permits are held by electricity installations, it is not surprising that many studies have found electricity prices to be a major determinant of carbon prices in the EU (e.g. Alberola et al, 2008b, Chevallier, 2009 and Hintermann, 2010).
- Temperature. The daily temperature of each EU country was obtained 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). Many studies have found temperature to be a significant price driver for carbon prices through its impacts on the demand for fuel (e.g. Christiansen et al, 2005, Mansanet-Bataller et al, 2007 and Hintermann, 2010).
- Availability of hydropower energy and precipitation levels. The first of these variables was constructed from data obtained from the Noordpool exchange, which is comprised 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 it by the total maximum capacity possible. Reservoir capacity will affect the supply of hydroelectric power as will the level of precipitation (e.g. Hintermann, 2010). The second variable is the precipitation level in two Noordpool countries for which complete daily data were available, Sweden and Finland.
- Stock price. Daily index of the stock prices of the most highly capitalized, 100 blue-chip companies in Europe (the Euronext 100). It was obtained from Thompson Datastream. Stock market returns are an indicator of the health of the economy. Economic growth is expected to raise the demand for energy and thus for carbon permits. However, we note that Chevallier

(2009), for example, has found no significant connection between carbon prices and stock prices.

- **Commodity price.** Daily index of world commodity prices (excluding energy). The index was obtained from Thompson Datastream. Chevallier (2009) includes a similar commodity price index, although it was not found to be a significant driver of carbon prices.
- **Corporate risk premium variable.** The daily risk premium is calculated as the difference between Moody's Aaa and Baa rated bonds. This is an indicator of the overall risk in the economy. Many asset pricing models (e.g. Fama and French, 1993), demonstrate how the risk premium varies with economic conditions. There will be less investment in a recession, when the risk premium is high. Conversely, when the economy is growing, the risk premium should be low and investors will expect a lower rate of return. By affecting investment, the risk premium will affect the demand for new facilities, employment and other measures of economic performance. Thus, it may have an impact on EUA prices.
- **Carbon prices in the US.** Daily carbon permit prices on the Chicago Climate Exchange (CCX). This is a voluntary carbon spot market in the US. As the largest carbon exchange in the US during the period covered in this study, it has the potential to affect the EU market, if a comprehensive mandatory cap-and-trade legislation were ever to be passed. For now, we surmise that a strong voluntary US market (strengthened by anticipation among traders of new mandatory legislation) may send a bullish signal to compliance markets elsewhere. The variable is measured in US dollars per metric tonne of carbon.
- **Carbon offset price.** We use the daily price of CER futures since 2008. We use the future rather than spot price 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. Since (in Phase 2) firms can cover a certain amount of their emissions with these carbon offsets, it is expected that lower offset prices will have a negative impact on the EUA price markets in Phase 2.
- **Overallocation of permits measure.** This variable is constructed by subtracting allocated permits from actual emissions each year. An oversupply of permits will be expected to have a depressive effect on prices. As noted, this was a particular problem in the first phase of the scheme.

It is also worth noting that our energy prices are one month ahead future prices (with one exception). The use of one month ahead futures is common in other studies (e.g. Hintermann, 2010 and Chevallier, 2009). As Alberola et al (2008a) note, EU installations under the scheme only need to hold permits at the end of each year and, thus, decisions relating to buying and selling of permits will likely reflect expectations of energy prices not on a specific day but

over the near future. One-month ahead energy prices provide the best indicator of the expectations of energy prices and are considered a less volatile market indicator than the intra-day spot market. The one exception is the McCloskey coal marker, which is only available on a weekly basis. It is considered by experts to be the best price information for hard coal for Europe (see Hintermann, 2010).

## 4 Statistical Methods

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

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 avoids 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 regressors 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. The reader is referred to Raftery et al (2010) for a detailed description of and additional motivation for DMA. Apart from the treatment of volatility (described below) our implementation of DMA is exactly as in Raftery et al (2010). The online appendix associated with this paper also provides complete statistical details.

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

$$\begin{aligned} y_t &= Z_{t-1}\theta_t + \varepsilon_t \\ \theta_{t+1} &= \theta_t + \eta_t, \end{aligned} \tag{1}$$

where  $y_t$  is the dependent variable,  $Z_{t-1}$  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 are always 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 and Primiceri, 2005). Standard methods for estimation and prediction exist with such models (see Koop and Korobilis, 2009, for a textbook level Bayesian treatment of TVP models).

They allow for the marginal effects of the predictors to change over time. However, TVP regression models can deal poorly with the sort of abrupt changes in the time series dynamics that we may expect with our data set. Furthermore, TVP regression models can be over-parameterized. That is, if  $m$  is moderately large (as in the present application), even a standard regression model can overfit the data. 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, Sala-i-Martin, Doppelhofer and Miller 2004, Eicher, Papa-georgiou and Raftery, 2010 and Ley and Steel, 2010). However, these papers use BMA with cross-sectional, homoskedastic, regression models. The contribution of Raftery et al (2010) was to develop computationally feasible methods of doing model averaging with TVP models such as (1). Empirical papers such as Dangl and Halling (2012) and Koop and Korobilis (2011) find DMA to be a useful approach when forecasting with financial or macroeconomic time series.

BMA with cross-sectional regressions involves averaging over regression models that differ in their explanatory variables. That is, let  $Z_t^{(k)}$  contain a subset of the explanatory variables in  $Z_t$  for  $k = 1, \dots, K$ . With  $m$  potential explanatory variables there are  $K = 2^m$  possible restricted versions of the unrestricted model that contains all the explanatory variables. BMA involves averaging over all these models. With TVP regressions, we have the same potentially large model space, but the added complication that the weights attributed to each model can vary over time. To be specific, with DMA, each of the models in the model space is:

$$\begin{aligned} y_t &= Z_{t-1}^{(k)} \theta_t^{(k)} + \varepsilon_t^{(k)} \\ \theta_{t+1}^{(k)} &= \theta_t^{(k)} + \eta_t^{(k)}, \end{aligned} \tag{2}$$

$\varepsilon_t^{(k)}$  is  $N(0, H_t^{(k)})$  and  $\eta_t^{(k)}$  is  $N(0, Q_t^{(k)})$ .

Let  $\pi_{t|t-1,k}$  be the probability that model  $k$  is the correct model for forecasting  $y_t$  given the data available at time  $t-1$ . DMA involves forecasting with all the models defined by (2). DMA forecasts are averaged across all these individual forecasts using  $\pi_{t|t-1,k}$  as weights. Dynamic model selection (DMS) can be done by forecasting  $y_t$  using only model  $j$  where  $\pi_{t|t-1,j}$  has the maximum model probability. DMA or DMS, thus, require forecasts from each individual model as well as calculation of  $\pi_{t|t-1,k}$  for  $k = 1, \dots, K$ . We describe these two requirements in turn.

Consider first forecasting within a single model. Given estimates of  $H_t^{(k)}$  and  $Q_t^{(k)}$ , this can be done in a straightforward fashion using the Kalman filter, which produces the predictive density of model  $k$  (see Raftery et al, 2010, equation 10). We denote this predictive density as  $p_k(y_t|y^{t-1})$  where  $y^{t-1} = (y_1, \dots, y_{t-1})'$ . To estimate  $Q_t^{(k)}$ , Raftery et al (2010) use a forgetting factor (sometimes called a discount factor) approach. Such approaches have long been used in the state space literature (e.g. West and Harrison, 1997). This specification implies that

observations  $j$  periods in the past have weight  $\lambda^j$  when estimating  $\theta_t$ . Following Raftery et al (2010), we set  $\lambda = 0.99$ , but consider other values as part of a sensitivity analysis. Note that, if  $\alpha = 0.99$ , an observation four weeks ago receives 80% as much weight in estimating  $\theta_t$  as the observation last period (when using daily data based on a five day week).

Raftery et al (2010) recommend setting  $H_t^{(k)} = H^{(k)}$  where  $H^{(k)}$  is a consistent estimate (see page 56). In our empirical section, we present results for homoskedastic versions of our models using this approach. However, when forecasting financial variables, it is likely that the error variance is changing over time. Thus, we use an Exponentially Weighted Moving Average (EWMA) estimate of  $H_t^{(k)}$ :

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

EWMA estimators are commonly used to model time-varying volatilities in finance; see, e.g., Riskmetrics (1996) or West and Harrison (1997) for the properties of EWMA estimators.  $\kappa$  is called a decay factor. Riskmetrics proposes setting  $\kappa = 0.94$  for daily data and our main results are for this choice (but we investigate sensitivity to it). An equivalent way of writing the EWMA estimator is (see West and Harrison, 1997, page 47):

$$\widehat{H}_t^{(k)} = \frac{(1 - \kappa)}{(1 - \kappa^t)} \sum_{j=1}^t \kappa^j \left( y_{t-j} - Z_{t-j-1}^{(k)} \widehat{\theta}_{t-j}^{(k)} \right)^2$$

from which it can be seen that past prediction errors are weighted in an exponential manner. For instance, if  $\kappa = 0.94$  then the prediction error four weeks ago receive 30% as much weight as last period's error.

Finally, we turn to the multi-model case and describe the calculation of  $\pi_{t|t-1,k}$  for  $k = 1, \dots, K$ . Raftery et al (2010) introduce another forgetting factor,  $\alpha$ , to do this. Their approach to DMA can be shown to imply:

$$\pi_{t|t-1,k} \propto \prod_{i=1}^{t-1} [p_k(y_{t-i}|y^{t-i-1})]^\alpha.$$

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,  $p_k(y_{t-i}|y^{t-i-1})$ ). 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 main value used in this paper and in Raftery et al, 2010), forecast performance four weeks ago receives 80% as much weight as forecast performance last period. Raftery et al (2010) show that the case  $\alpha = \lambda = 1$  is equivalent to standard BMA done in a constant coefficient model using marginal likelihoods as weights in the model averaging. We consider this special case in our empirical work.

In summary, the DMA approach used in this paper requires the choice of three key factors:  $\lambda, \kappa$  and  $\alpha$ , that control the degree of variation in coefficients,

volatilities and model switching, respectively. The main results below set  $\lambda = 0.99$ ,  $\kappa = 0.94$  and  $\alpha = 0.99$ , although we investigate the sensitivity to these choices. Given these forgetting factors, DMA involves only the use of recursive filtering algorithms for state space models including the Kalman filter and one derived in Raftery et al (2010).

In terms of the properties of DMA or DMS, the advantage is that they allow for switches between parsimonious models. That is, instead of suffering the overfitting 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 coefficient models or models with too few price drivers are used), DMS can switch, for example, from 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 and DMS have been found to favour parsimonious models and can handle abrupt changes in the dynamic structure much better than conventional TVP models.

## 5 Forecasting Carbon Prices in the EU ETS

### 5.1 Introduction

All of our models include an intercept and a lag of the dependent variable. Our models differ 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 and DMS is selecting the best of these 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 (which are the precipitation and over-allocation of permits variables). 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 are discussed. For our discussion of price drivers, we present a selection of results for the full sample of the futures data. Complete results are available in the online appendix.

### 5.2 Price Drivers

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 the entire data set is used. Remember that, when forecasting  $y_t$ , DMA is using  $\pi_{t|t-1,k}$  for each of  $k = 1, \dots, K$  models as weights in the model averaging exercise. We can use  $\pi_{t|t-1,k}$  in various ways to shed light on which (and how many) price drivers DMA is using at each point in time. Figure 2 sheds light on how parsimonious DMA is. Let  $Size_{(k)}$  be the number of price drivers in model  $k$  then

$$E(Size_t) = \sum_{k=1}^K \pi_{t|t-1,k} Size_{(k)}$$

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.

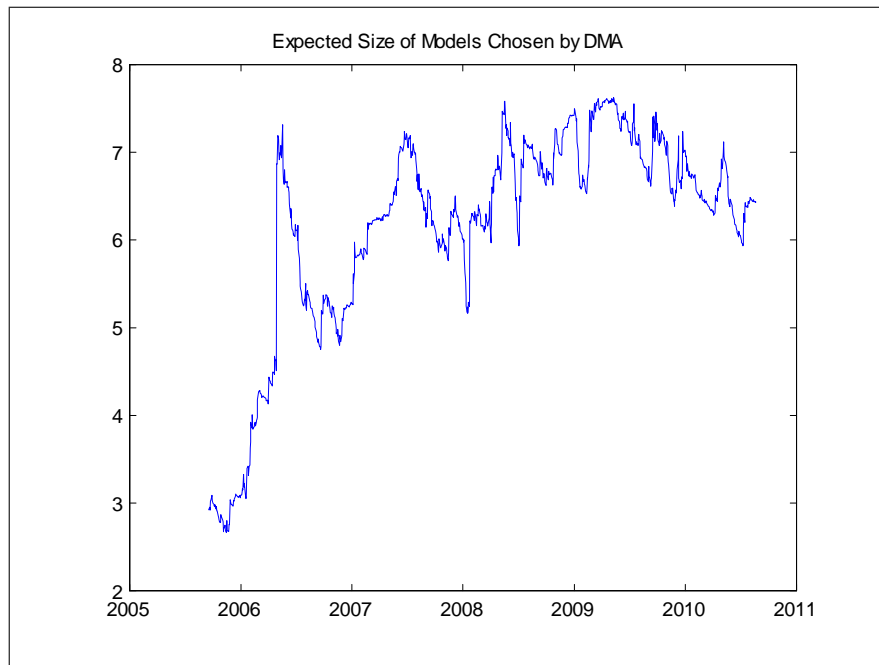


Figure 2: Results for Carbon Futures Data

We can use  $\pi_{t|t-1,k}$  to provide evidence on which price drivers are receiving the most weight in the model averaging exercise by calculating posterior inclusion probabilities. The inclusion probability of the  $j^{th}$  price driver is defined as the probability DMA attaches to models that include the  $j^{th}$  price driver. Figures 3 through 5 plot these inclusion probabilities. They show that constant coefficient models that simply include all the price drivers (or a selected few) are inappropriate since there is a great deal of variation over time in respect to which price drivers are included.

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.

In terms of financial variables, the corporate risk premium is sometimes important and briefly becomes very important at the height of the financial crisis in the autumn of 2008. In terms of the variables reflecting climate, they fluctuate in importance and timing. For instance, the temperature variable becomes briefly important in the latter half of 2006, but the precipitation variable becomes important in the latter half of 2007.

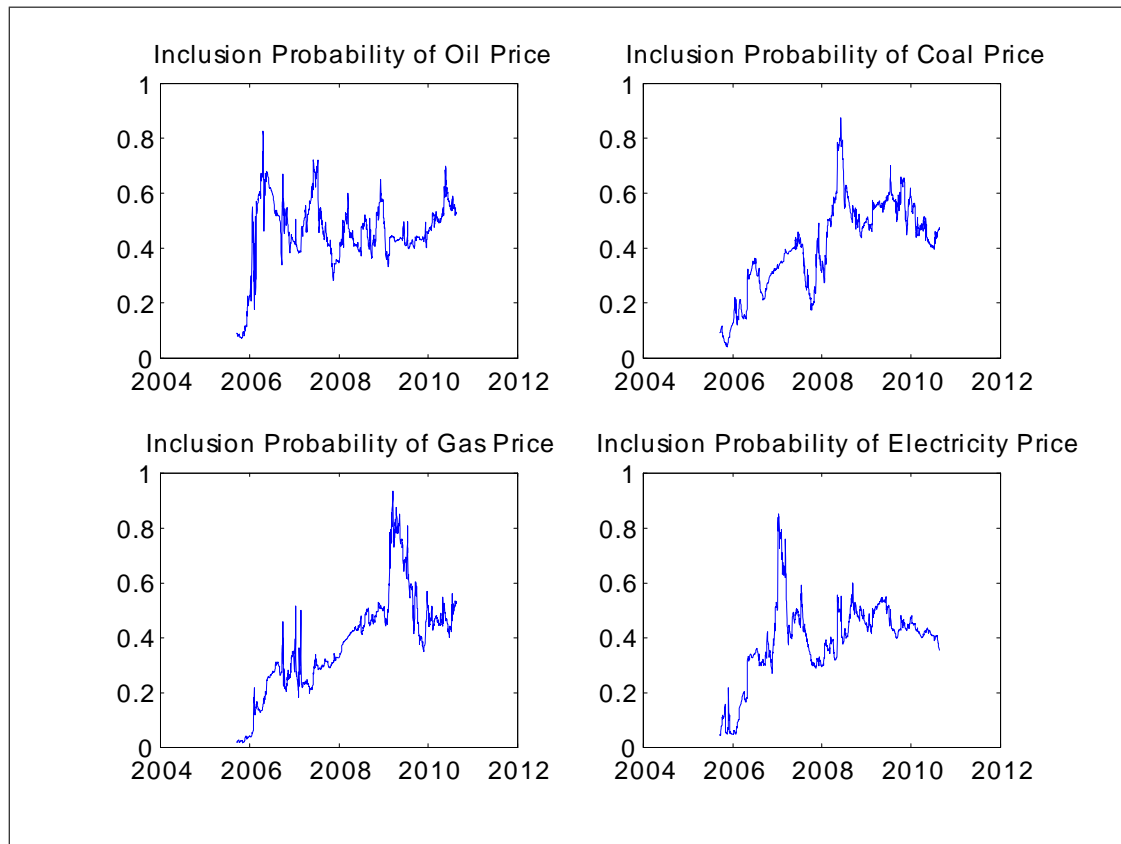


Figure 3: Results for Carbon Futures Data



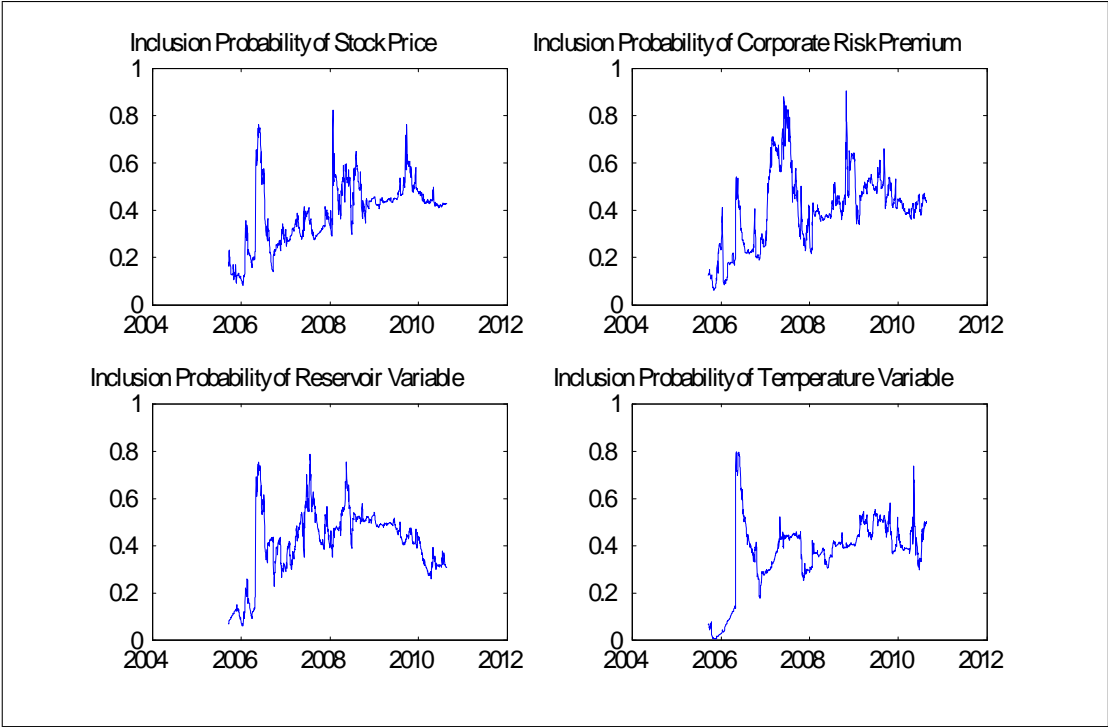


Figure 4: Results for Carbon Futures Data

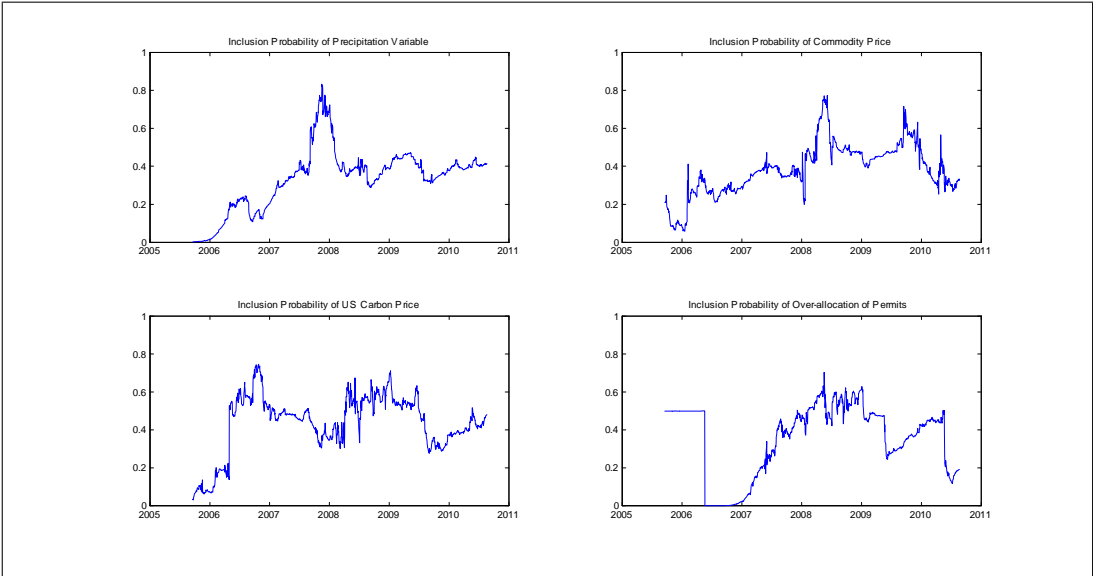


Figure 5: Results for Carbon Futures Data

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 using for spot prices and using only Phase 2 data. The online appendix also plots the error variances. Suffice it to note here 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 and DMS, 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 is a measure of 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. 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 dynamic model selection (DMS) to several alternatives with 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 and error variances is removed (i.e. as described in Section 4, it implements BMA as used with cross-sectional regression models and is obtained using a homoskedastic version of the model with  $\alpha = \lambda = 1$ ). Otherwise, all the assumptions are the same as in DMA. Similarly, BMS is the same as DMS but in a constant coefficient model.

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 carrying out the DMA but does no model averaging.

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.

In order to investigate the sensitivity of our results to choices of forgetting factors, we also present versions of DMA and DMS with other values for these factors and list these values as arguments for the acronyms. For instance,  $\text{DMS}(\alpha = 0.99, \lambda = 0.95)$  is the same as DMS except we set  $\lambda = 0.95$ , thus allowing for a greater degree of change in the coefficients.

Finally, we investigate the importance of heteroskedasticity through the use of different values for the decay factor,  $\kappa$ . Thus,  $\text{DMA}(\kappa = 0.97)$  is the same as DMA (and both have  $\alpha = \lambda = 0.99$ ) except less variation in volatility is allowed for.  $\text{DMA}(\text{homo})$  is the homoskedastic version of DMA.

Table 1 presents the sums of log predictive likelihoods for these different approaches. Our main finding is that DMA or DMS do tend to forecast better than the other approaches and, in many cases, the forecast improvements are substantial. This statement holds true for spot or future prices and for both the full sample and Phase 2.

The TVP regression model forecasts poorly throughout, due to its over-parameterization. However, simply removing the time-variation in coefficients and working with a constant coefficient model does not lead to better forecasts. CC-NONINF and SHRINK also forecast poorly (although, when comparing these two forecasting methods, it is worth noting that prior shrinkage does lead to some improvements in forecast performance). Similarly, conventional BMA in a constant coefficient model does not forecast well.

Our benchmark choice of  $\kappa = 0.94$  means that we allow for a substantial amount of change in the error variance over time. Allowing for less variation in the error variance leads to a worsening of forecast performance and, in particular, the homoskedastic versions of DMA and DMS forecast poorly.

Our choice of forgetting factors was based on the recommendations of Raftery et al (2010). The results in Table 1 suggest that these recommendations are reasonable, but that results are fairly robust to changes in them. With Phase 2 data, the best forecasts of both spots and futures are obtained with the benchmark choices of  $\alpha = \lambda = 0.99$ . With the full sample of spot prices, allowing for more rapid model change leads to slight improvements in performance. That is, DMA with  $\alpha = 0.95, \lambda = 0.99$  has the highest sum of log predictive likelihoods. For the full sample of future prices, DMS with  $\alpha = \lambda = 0.95$  has the best forecast performance, indicating that allowing for more rapid model and coefficient change can lead to some slight forecast improvements relative to our benchmark choices.

There do not seem to be any systematic differences between model averaging and model selection approaches. There clearly are large benefits in allowing for model change; approaches that do forecast better than those that do not. However, model selection and model averaging approaches forecast similarly and there is no consistent pattern found in which one approach appears better than the other.

Table 1 is based on sums of log predictive likelihoods over the entire sample (apart from the initial 100 observations). Graphs of the cumulative sums up to each point in time are available in our online appendix. 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), the graph 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.

	Spot	Futures	Spot (Phase 2)	Future (Phase 2)
DMA	1,354	2,831	1,286	1,307
DMS	1,347	2,822	1,289	1,308
BMA	1,235	2,401	1,014	1,024
BMS	1,245	2,350	1,014	1,025
TVP	681	2,773	1,239	1,245
CC-NONINF	409	2,803	1,203	1,266
SHRINK	461	2,813	1,233	1,284
DMA( $\alpha = \lambda = 0.95$ )	1,259	2,832	1,234	1,260
DMS( $\alpha = \lambda = 0.95$ )	1,286	2,842	1,259	1,287
DMA( $\alpha = 0.95, \lambda = 0.99$ )	1,367	2,829	1,277	1,295
DMS( $\alpha = 0.95, \lambda = 0.99$ )	1,349	2,808	1,277	1,307
DMA( $\alpha = 0.99, \lambda = 0.95$ )	1,240	2,801	1,271	1,294
DMS( $\alpha = 0.99, \lambda = 0.95$ )	1,249	2,811	1,283	1,301
DMA( $\kappa = 0.97$ )	1,301	2,770	1,200	1,219
DMS( $\kappa = 0.97$ )	1,311	2,774	1,205	1,224
DMA(homo)	1,287	2,366	997	1,007
DMS(homo)	1,270	2,375	1,011	1,021

Table 2 is in the same format as Table 1, but presents MSFEs instead of sums of log predictive likelihoods. Overall, Table 2 tells the same story as Table 1. DMA and DMS do lead to substantial improvements in forecast performance over regression methods, even if the regression methods allow for time-varying coefficients. However, there is one substantive difference between Tables 1 and 2 that is worth noting. Forecasting methods with homoskedastic errors tend to perform much better when the metric is MSFE than when the metric is the sum of log predictive likelihoods. In fact, the homoskedastic version of DMS often has the lowest MSFE of any of the approaches we have considered. This finding shows the role of the error variance in forecasting. A good estimate of the error variance is not always necessary to obtain a good point forecast. Hence MSFEs can be low for the homoskedastic model even if heteroskedasticity is present. But Table 1 tells us that appropriate modelling of the error variance is important in obtaining a good estimate of the entire predictive distribution.

Table 2: Mean Squared Forecast Errors

	Spot	Futures	Spot (Phase 2)	Future (Phase 2)
DMA	60.66	0.942	0.413	0.381
DMS	61.55	0.978	0.426	0.390
BMA	59.01	0.929	0.414	0.386
BMS	56.17	0.939	0.417	0.393
TVP	122.98	1.026	0.449	0.417
CC-NONINF	271.66	0.984	0.464	0.408
SHRINK	233.55	0.975	0.447	0.406
DMA( $\alpha = \lambda = 0.95$ )	109.26	0.987	0.458	0.424
DMS( $\alpha = \lambda = 0.95$ )	163.87	0.984	0.493	0.451
DMA( $\alpha = 0.95, \lambda = 0.99$ )	59.28	0.941	0.417	0.382
DMS( $\alpha = 0.95, \lambda = 0.99$ )	60.89	0.968	0.440	0.399
DMA( $\alpha = 0.99, \lambda = 0.95$ )	141.87	1.043	0.431	0.398
DMS( $\alpha = 0.99, \lambda = 0.95$ )	239.23	1.039	0.440	0.404
DMA( $\kappa = 0.97$ )	58.23	0.945	0.416	0.382
DMS( $\kappa = 0.97$ )	56.28	0.987	0.433	0.397
DMA(homo)	57.60	0.932	0.407	0.374
DMS(homo)	56.03	0.931	0.404	0.373

#### 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 on 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 are still producing substantial improvements in forecast performance in Phase 2 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 the 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 market could have an impact on emitters, making it more difficult for them to plan ahead and achieve reductions in CO<sub>2</sub> emissions in a cost effective manner.

## 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 in the EU ETS market. We show how our DMA approach can model these features and forecast accurately compared to other approaches.

## 7 References

- Alberola, E., Chevallier, J., and Cheze, B. (2008a) The EU emissions trading scheme: disentangling the effects of industrial production and CO<sub>2</sub> emissions on carbon prices. *Int. Econ.* 116, 93-126.
- Alberola, E., Chevallier, J., and Chèze, B. (2008b) Price drivers and structural breaks in European carbon prices 2005–2007. *Energy Pol.* 36, 787–797.
- Alberola, E., Chevallier, J., and Chèze, B. (2009a) Emissions compliances and carbon prices under the EU ETS: A country specific analysis of industrial sectors. *J. Pol. Model.*, 31, 446-462.
- Alberola, E., Chevallier, J. and Chèze, B. (2009b) European carbon prices and banking restrictions: evidence from Phase 1 (2005-2007). *Energy J.* 30, 51-80.
- Avramov, D. (2002) Stock return predictability and model uncertainty. *J. Finan. Econ.* 64 423–458.
- Bredin, D. and Muckley, C. (2011) An emerging equilibrium in the EU emissions trading scheme. *Energy Econ.* 33, 353-362.
- Bunn, D. and Fezzi, C. (2007) Interaction of European carbon trading and energy prices, *Fondazione Eni Enrico Mattei Working Papers*, Paper 123.
- Chevallier, J. (2009) Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Econ.* 31, 614-625.
- Christiansen, A., Arvanitakis, A., Tangen, K., and Hasselknippe, H., (2005) Price determinants in the EU emissions trading scheme. *Clim. Pol.* 5, 15-30.
- Cogley T. and Sargent T. (2005) Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S. *Rev. Econ. Dyn.*, 8, 262-302.

Cogley T., Morozov S. and Sargent T. (2005) Bayesian fan charts for U.K. inflation: Forecasting and sources of uncertainty in an evolving monetary system. *J. Econ. Dyn. Control* 29, 1893-1925.

Convery, F., Redmond, L. (2007) Market and price developments in the European Union Emissions Trading Scheme. *Rev. Environ. Econ. Policy* 1, 88–111.

Dangl, T. and Halling, M. (2012) Predictive regressions with time varying coefficients. *J. Fin. Econom.* forthcoming.

Eicher, T., Papageorgiou, C. and Raftery, A. (2010) Determining growth determinants: Default priors and predictive performance in Bayesian model averaging, with applications to growth determinants. *J. Appl. Econom.*, forthcoming.

Fama, E. F. and French, K. R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *J. Fin. Econom.* 25, 3-56.

Fernandez, C., Ley, E. and Steel, M. (2001a) Model uncertainty in cross-country growth regressions. *J. Appl. Econom.* 16, 563-576.

Fernandez, C., Ley, E. and Steel, M. (2001b) Benchmark priors for Bayesian mode averaging. *J. Econom.* 100, 381-427.

Frunza, M.-C., Guegan, D. and Thiebaut, D. (2010) Missing trader fraud on the emissions market, *Documents de Travail du Centre d'Economie de la Sorbonne*, 2010.71.

Hintermann, B. (2010) Allowance price drivers in the first phase of the EU ETS. *J. Environ. Econ. Manag.* 59, 43-56.

IETA (International Emissions Trading Association) (2010). GHG Market Report 2010. Available through the website <http://www.ieta.org>.

Kanen, J.L.M. (2006). *Carbon Trading and Pricing*. Environmental Finance Publications: London, UK.

Klein Tank, A.M.G., Wijngaard, J., Können, G., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L. and Petrovic, P. (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Internat J. Climatol.* 22, 1441-1453. Data and metadata available at <http://eca.knmi.nl>.

Koop, G. and Korobilis, D. (2009) Bayesian multivariate time series methods for empirical macroeconomics. *Found. Trends. Econom.* 3, 267-358.

Koop G and Korobilis, D. (2011) Forecasting inflation using dynamic model averaging. *Int. Econ. Rev.*, forthcoming.

Kossoy, A. and Ambrosi, P. (2010) *State and Trends of the Carbon Market 2010*. World Bank: Washington, D.C.

Ley, E. and Steel, M. (2009) On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. *J. Appl. Econom.* 24, 651-674.

Ley, E. and Steel, M. (2010) Mixtures of g-priors for Bayesian model averaging with economic applications, Centre for Research in Statistical Methodology, Working Paper No. 10-23.

Mansanet-Bataller, M., Chevallier, J., Hervé-Mignucci, M. and Alberola, E. (2010) The EUA-sCER Spread: Compliance strategies and arbitrage in the European carbon market, Mission Climat Working Paper, 2010-6. Available through the website <http://www.caissedesdepots.fr>.

Primiceri G. (2005) Time varying structural vector autoregressions and monetary policy. *Rev. Econ. Stud.* 72, 821-852.

Raftery, A., Karny, M. and Ettler, P. (2010) Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill, *Technom.*, 52, 52-66.

Raftery, A., Madigan, D. and Hoeting, J. (1997) Bayesian model averaging for linear regression models. *J. Amer.Stat. Assoc.* 92, 179-191.

RiskMetrics. (1996) Technical Document (Fourth Edition) available at <http://www.riskmetrics.com/system/files/private/td4e.pdf>.

Sala-i-Martin, X., Doppelhofer, G. and Miller, R. (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. *Amer. Econ. Rev.* 94, 813-835.

West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models (second edition)*. Springer: New York.