State-space risk measurement: An application to renewable energy returns

Julian Inchauspe

Centre for Research in Energy and Minerals Economics, Curtin Business School, Curtin University, Perth, Australia
Email: Julian.Inchauspe@curtin.mq.edu.au

Abstract

This paper uses state-space methodology for modelling excess returns, risk and dynamics for the WilderHill New Energy Index (NEX). The NEX is a global exchange-traded index for investment in development, production and efficiency of renewable energy. It currently lists 98 companies located in 21 countries; the total capitalization of the index is about 285 billion US$ (www.nexindex.com). The NEX has experienced a substantial growth in the last decade along with the rapid development of the renewable energy sector. According to UNEP (2010) estimations, the total amount of (public and private) new investment in renewable energy increased from 46 billion US$ in 2004 to about 162 billion US$ in 2009. As a result, renewable power generation capacity has increased from about 4% of total power generation to nearly 7% between these two dates. Along with this long-run positive trend, the NEX has been able to offer high returns. Naturally, these returns have been associated with high risk exposure. For instance, the index suffered substantial turbulence between 2007 and the end 2009. This paper is aimed at bringing a deeper understanding of the fundamentals that underpin this behaviour. The models in this paper considers various fundamentals that have been associated with the NEX in various reports. The analysis is carried with weekly data between week 33 in 2001 and week 12 in 2011.

This study reports work in progress on two different state-space methodologies for assessing the returns and risk for the NEX. First, I use a multi-factor state-space model with time-varying coefficients to analyze the impact of different fundamental variables on the NEX. This first approach was applied to monthly data in Inchauspe, Ripple & Trueck (2011) and presented at the 34th International Conference by the International Association for Energy Economics. Expanding the research to weekly data suggests problems in the robustness of that model. To remedy this, I propose using a Markov-switching (MS) model that allows for regime inference and dynamic analysis. The MS model is written in a mean-adjusted form. This mean, as well as the covariance matrix, are allowed to shifts over time. This specification allows for assessing the significance of exogenous variables after allowing for shifts in mean NEX returns. As regime shifts in NEX excess returns are associated with a positive trend in the NEX levels, the regimes are labelled as “bull” or “bear” markets that cannot be explained by fundamentals.

Earlier literature has proposed using state-space methodology to measure time-varying beta factors in capital asset pricing specification (e.g. Bolleshev, Engle and Wooldridge, 1988; Koopman et al., 2008; van Geloven and Koopman, 2009; Tsay, 2005, p.577). The first model borrows from this approach to specify a multi-factor model with time-varying coefficients. In addition, a considerable amount of literature has used Markov-switching models to study univariate dynamics of “bull” and “bear” markets in stock market indices (Gordon and St-Amour, 2000; Maheu and Curdy, 2000; Lunde & Timmerman, 2004, Edwards et al., 2003; Girardin and Liu, 2003; Pagan & Soussonov, 2003). I propose studying the dynamics of possible bull/bear markets after relevant exogenous fundamentals are included in the analysis; this approach has also been popular in the literature (Chen, 2009; Chang, 2009; Guidolin & Timmermann, 2005; Chauvet & Porter, 2001). The Markov-switching model in the second part of this paper allows for distinguishing four distinctive distributions associated with abnormal returns and variances. The information obtained from this estimation is valuable for analysts and investors considering medium-run long positions and provides insights into mean-reverting properties of NEX excess returns.

Keywords: State-Space Modelling; Markov-Switching; Risk; Renewable Energy.
1. INTRODUCTION

The renewable energy sector has accomplished a remarkable progress at global level during the last decade. At global level, it has been estimated that private and public investment in renewable energy climbed from 46 billion US$ in 2004 to 162 billion in 2009 (UNEP, 2010). Despite this positive trend, these estimates also suggest a slowdown in the 2009/2010 period.

In this paper, we study excess returns for the WilderHill New Energy Global Index (NEX), which is subset of total investment. The NEX plays an important role in the valuation of renewable energy assets in the global market. Its total capitalization is about 285 billion US$ (www.nexindex.com); if we consider that the last available estimator for public and private investment in renewable energy is 162 billion US$, it becomes clear the NEX plays an important role in the renewable energy sector. UNEP reports on renewable energy often dedicate some space to the analysis of the NEX (UNEP, 2010). The NEX has evolved rapidly in the last decade, following the long run trend on renewable energy. However, the index behaves significantly different to the aggregate estimates in the short and medium run. UNEP’s Global Trends in Sustainable Energy Investment report (UNEP, 2010, p.30) and the NEX official webpage (www.nexindex.com) compare the evolution of the NEX to other leading indices such as Amex Oil, Nasdaq composite and S&P 500.

Previous literature investigating the returns of renewable energy indices has focused on the influence of these factors. Some studies have focused on the WilderHill Clean Energy Index (ECO). Compared to the NEX, the ECO has a smaller capitalization and includes US firms mostly, but is similar in all other aspects; in fact, there is a strong degree of correlation between the two. A study using a vector autoregression for returns on the ECO, WTI oil prices, PSE (technology) index and US interest rate has suggested that the ECO returns are highly influenced by the PSE technology index and that oil prices are not influential (Henriques and Sadorsky, 2008). A later article suggested that oil prices contribute to ECO systematic risk (Sadorsky, 2010). Another study based on vector autoregressive methodology considers the ECO and different measures of oil price –WTI price, Brent price, futures price and an oil index– concluding that the period 2001-2010 is better explained if it is split in three sub-periods (Chen et al., 2011). The price of the NEX and the ECO have also been studied with a vector autoregression including a technology index, the S&P 500 index, oil price and the European Emission Trading carbon price (Kumar et al., 2011). The disadvantages of the latter are that the European carbon price used in this estimations may not be a good representation of the carbon price in other parts of the world, and that the variance decomposition imposes strong assumptions into the model. Inchauspe, Ripple & Trueck (2011) have analyzed the influence of some factors in a time-varying multifactor setup. To the best of our knowledge, there are no other major contributions studying the WilderHill indices.

Overall, these studies suggest that technology indices and benchmark stocks have strong influence on the renewable energy indices while the influence of oil is weak. It has been suggested that investors may view renewable/alternative energy companies as similar to high technology stocks (Henriques and Sadorsky, 2008). The latter has strong implications for the investor. For instance, investment in oil is sometimes used to hedge risk in a diversified portfolio while, on the other hand, high technology stocks tend to yield high returns and be very volatile. This observation is one of the motivating ideas for this work.

This paper studies different aspects of the NEX in two sections. The first part of the paper uses a Kalman-filter multifactor model which allows all beta factors to become time-varying. State-space models with time-varying beta factors have been previously used in financial literature (e.g. Bolleshev, Engle and Wooldridge, 1988; Koopman et al., 2008; van Geloven and Koopman, 2009; Tsay, 2005, p.577). The estimations suggest that, with weekly data, the above model brings certain disadvantages (in Inchauspe et al., 2011, monthly data was used). This question motivates the second part of the paper. Seeking a model which could fit and explain the weekly data, I propose a Markov-switching autoregressive distributed lag model to study the dynamics and distributional properties of “bull” and “bear” markets after fundamental and autoregressive components are taken into account. The identification of bear/bull markets in univariate analysis with Markov switching models has been popular in the literature (Gordon and St-Amour, 2000; Maheu and Curdy, 2000; Lunde & Timmerman, 2004, Edwards et al., 2003; Girardin and Liu, 2003; Pagan & Soussonov, 2003), and some authors have suggested using exogenous variables (Chen, 2009; Chang, 2009; Guidolin & Timmermann, 2005; Chauvet & Porter, 2001). The Markov-switching approach also allows for medium run analysis of returns and variance, as will be shown later.

In the dataset, several exogenous variables are considered. First, excess returns on the WTI oil price are included; we think that the global oil price may play an important role in the development of renewable energy and that its recent spiky behaviour around 2008 may have been influential. Second, we also consider the Amex Oil index as an explanatory variable. The Amex Oil index is linked to the exploration, development and production of oil; all these activities compete with the renewable energy sector for the same
resources. Third, returns for the Nasdaq composite index are considered, following the suggestion of the literature that technology stocks may be influential (Henriques and Sadorsy, 2008). The literature cited above has also suggested that the interest rate plays an important role in determining NEX returns, and most likely in determining the above variables too. To deal with this problem, I use excess returns for all variables. For the weekly dataset in this paper, excess returns are calculated by subtracting from the weekly returns the weekly equivalent interest rate of a 4-week US treasury bond (this interest rate is interpreted as the opportunity cost of a risk-free short-term investment). The dataset includes weekly observations between the first week of December 2001 and the second week of June 2011.

The balance of this paper is organized as follows. Section 2 is dedicated to the first model. A multifactor model with time-varying coefficients is proposed and estimation results are discussed. In Section 3, a Markov-switching mean-adjusted autoregressive distributed lag model is estimated. Section 4 concludes the paper.

2. A MULTIFACTOR MODEL WITH TIME-VARYING PARAMETERS

The state-space model proposed in this section has a similar structure to other models used in the literature (van Geloven and Koopman, 2009; Tsay, 2005, p.577, Koopman et al., 2008; Kim and Nelson, 1999, p.44; and Bolerslev, Engle and Wooldridge, 1988), although none of these authors applied this methodology to energy company equities. The model can be defined as follows:

\[ r_{\text{NEX},t} = \alpha_t + \beta_{\text{SP}} r_{\text{SP},t} + \beta_{\text{PSE}} r_{\text{PSE},t} + \beta_{\text{OIL}} r_{\text{OIL},t} + \epsilon_t, \quad \epsilon_t \sim \text{nid}(0, \sigma^2_t), \]  

(1)

The novel aspect of the proposed model is the introduction of time-varying coefficients \( \alpha, \beta_{\text{SP}}, \beta_{\text{PSE}} \) and \( \beta_{\text{OIL}} \) that follow a random-walk specification updated through a Kalman-filter. The evolution of the time-varying coefficient is reported in Figure 1 and the out-of-sample forecast is reported in Figure 2. The Amex Oil and Nasdaq indices are, overall, strongly significant factors (with \( p \)-values of 0.0006 and 0.0000 respectively). The constant (i.e. the abnormal returns) and S&P500 excess returns are not significantly different than zero, while oil price returns are only weakly significant (\( p \)-value: 0.1113). These results have to be taken with caution and compared with the dynamic evolution of the time-varying coefficients in Figure 1. The latter suggests that the beta factor for the S&P 500 might be significant in some subperiods while the beta coefficient for the oil price remains smooth around zero. The beta factor for Amex Oil increases up to 2008 which is consistent with the NEX market becoming better related to its fundamentals as it grows. As pointed earlier, the Amex Oil index list firms linked to geological research, exploration and development of oil; the firms that carry these activities compete for human and other type of resources with the renewable energy sector. The influence of the Nasdaq index peaks in 2005-6 and decline afterwards. The out-of-sample one-period ahead forecast in this model is based on the previous period Kalman-filter estimates (without smoothing). While this model provides some insightful information, it is also subject to some inconveniences in terms of its covariance structure. With time-varying coefficients, one would expect enough flexibility as to expected well-distributed residuals with no volatility clusters. However, after running a GARCH(1,1) model on the irregular disturbance \( \epsilon_t \), I find evidence in favour of a volatility clusters that become accentuated between late 2008 until beginnings of 2010 (Figure 3). This finding would partially invalidate the identification of risk by simple standardization of the time-varying estimates. Adapting the proposed model to include conditional heteroskedasticity is both trivial and difficult, if the two-step procedure is to be avoided (conditional heteroskedasticity interferes with the Kalman-filter assumptions). Facing these difficulties, I decide not to elaborate any further on this model and employ another state-space model that will allow us to overcome the issue while providing new insights.
In this section, I use a MSMH-ADL (Markov-switching mean-adjusted heteroskedastic autoregressive distributed lag) model. In an ADL model, the mean-adjusted dependent variable is expressed as a function of its own past and current and lagged values of exogenous variables. The mean of the dependent variable is allowed to switch according to first-order, homogenous Markov process whose parameters are inferred in an expectation maximization algorithm. The variance of the model is also allowed to switch. A switch in NEX returns implies a change of trend in the NEX levels. Based on this approach, I identified abnormal bull (bear) markets which are defined as periods of generally increasing (decreasing) NEX prices which cannot be explained by its past or by exogenous variables. More formally, a MSMH-ADL model is written as:

$$[R_t - \mu_R(s_t)] = \sum_{i=1}^k \phi_i [R_{t-i} - \mu_R(s_{t-i})] + \sum_{q=1}^n \beta_q (X_q - \mu_X) + \epsilon_t, ~ \epsilon_t \sim N(0, \sigma(s_t)).$$

Where $s_t = j$ indicates which of the $[j, ..., M]$ regimes prevail in the system at time $t$. Under this model assumptions, the probability of jumping from a state $i$ to a state $j$ is expressed as $p_{ij} = Pr[s_t = j|s_{t-1} = i]$ and collected in the following transition matrix:

$$P = \begin{pmatrix}
    p_{11} & p_{12} & \cdots & p_{1M} \\
    p_{21} & p_{22} & \cdots & p_{2M} \\
    \vdots & \vdots & \ddots & \vdots \\
    p_{M1} & p_{M2} & \cdots & p_{MM}
\end{pmatrix}, \sum_{i=1}^M p_{ij} = 1.$$

The value of the elements in this transition matrix is unknown and unobservable. These values and the regime classification are inferred in a statistically efficient way in the estimation procedure. Unlike other studies (such as Pagan and Sossounov, 2000) we do not impose minimum duration constraints for the regimes. Compared to the previous state-space model, the estimation algorithm presents two major differences. First, the Kalman filter cannot be used and a special filter and smoother is used instead. Second, to overcome conditional definitions in the density function, the estimation is carried via an iterative Expectation Maximization (EM) algorithm (a full description of the estimation procedure is given in Kim and Nelson, 1999, Ch. 5; the latter is similar to the procedure described in Hamilton, 1994, Ch. 22). The estimators obtained with this method are asymptotically efficient (see Krolzig, 1997).

For the 4 regimes in this model, we use the following labels: Regime 1- Highly speculative ‘bear’ market; Regime 2- ‘bear’ market; Regime 3- ‘bull’ market; Regime 4- Highly speculative ‘bull’ market. The term ‘highly speculative’ simple mean that these regimes are expected to yield higher returns and be associated

\[1 \text{ The main difference is that Kim & Nelson (1999) use smoothing techniques that are easier to compute than Hamilton (1989, 1994).}\]
with a higher variance compared to the ‘moderate’ bull and bear market regimes. A priori, our expectations about means and volatilities in a $M=4$ setup are as follows:

\[
\mu(1) < \mu(2) < \mu(3) < \mu(4); \mu(1), \mu(2) < 0; \mu(3), \mu(4) > 0; \sigma(1) > \sigma(2); \sigma(4) > \sigma(3). \quad (8)
\]

The final estimation results are summarized in Table 1. For selecting the structure of the model, likelihood-ratio (LR) tests, Akaike’s information criterion (AIC) and t-ratios were used. The coefficients associated with WTI oil price excess returns were found not statistically different than zero in all the specifications that were considered, so it was excluded from the analysis. The Amex Oil and Nasdaq excess returns are highly significant, while the S&P 500 excess returns are marginally significant. The autoregressive coefficients for NEX were not significant in this model, and with a simple linear autoregression I could verified that they were not significant in this alternative model either. The proposed model produces significant improvement with respect to its linear counterpart according to the LR test in Table 1.

<table>
<thead>
<tr>
<th>Estimation Report</th>
<th>Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>Weekly</td>
</tr>
<tr>
<td>Number of autoregressive lags</td>
<td>0</td>
</tr>
<tr>
<td>Number of Markov-switching regimes</td>
<td>4</td>
</tr>
<tr>
<td>Number of parameters in linear model</td>
<td>24</td>
</tr>
<tr>
<td>Number of autoregressive lags</td>
<td>0</td>
</tr>
<tr>
<td>Number of Markov-switching regimes</td>
<td>4</td>
</tr>
<tr>
<td>Number of parameters in linear model</td>
<td>24</td>
</tr>
<tr>
<td>Number of parameters in linear model</td>
<td>6</td>
</tr>
<tr>
<td>Estimation Output vs. Linear Model</td>
<td>MSMHAR Linear</td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>1310.93 1214.18</td>
</tr>
<tr>
<td>AIC</td>
<td>-5.0567  -4.7413</td>
</tr>
<tr>
<td>Linearity Test</td>
<td>193.49</td>
</tr>
<tr>
<td>LR Chi-Square Statistic</td>
<td>0.0016</td>
</tr>
<tr>
<td>Markovian Dynamics</td>
<td>Number of Obs.  Ergodic Prob. Av. Duration</td>
</tr>
<tr>
<td>Regime 1: Highly Speculative Bear Market</td>
<td>62.7 0.0994 7.35</td>
</tr>
<tr>
<td>Regime 2: Moderate Bear Market</td>
<td>215.4 0.5375 149.93</td>
</tr>
<tr>
<td>Regime 3: Moderate Bull Market</td>
<td>89.0 0.1441 6.20</td>
</tr>
<tr>
<td>Regime 4: Highly Speculative Bull Market</td>
<td>141.9 0.2220 6.75</td>
</tr>
<tr>
<td>Transition Probability Matrix</td>
<td></td>
</tr>
<tr>
<td>$P_{11}$</td>
<td>0.050787</td>
</tr>
<tr>
<td>$P_{21}$</td>
<td>0.014276</td>
</tr>
<tr>
<td>$P_{31}$</td>
<td>0.006007</td>
</tr>
<tr>
<td>$P_{41}$</td>
<td>0.022599</td>
</tr>
</tbody>
</table>

Table 1 - Estimation summary: MSMH-ADL model.

Our results are in line with our expectations from equation (8). Highly-speculative bear (bull) markets regimes have higher absolute mean and variance than the moderate bear (bull) markets. Regime 2 (moderate bear market) is the most persistent, with an average duration of about 215 weeks. Regime 4 is the most recurrent state (it can be accessed from other states more easily than the other regimes). Conditional on the regime classification, NEX excess returns are positively correlated to those of Amex, Nasdaq and S&P 500. A full picture of the regime classification is provided in Figure 4. Regime 2 dominates between 2003 and beginnings of 2005 and from 2009, which is consistent with a mean-reverting property in the data. Abnormal moderate- and speculative-bull regimes dominate between 2005 and 2008. The highly-speculative bear regime is present during the start of the 2008-9 global financial crisis. It is worth remarking that although the exogenous variables were also affected by global financial crisis, what is measured with $\mu(5_i)$ is the excess reaction of the NEX. Furthermore, the residuals show no serial autocorrelation with the coefficients in a GARCH(1,1) specification being not significantly different than zero. This model is not subject to the volatility cluster limitation encountered earlier.

Using the information from the transition matrix, we can reconstruct long-term probabilities which should converge to the ergodic values $P = \lim_{n \to \infty} P^n$. This has been done in Figure 5 which suggests that unconditional (i.e. not depending from the initial state) ergodic probabilities are achieved after approximately 140 weeks. Based on this information, Table 2 provides a tool for investors who consider taking long positions over 140 weeks in a portfolio. These results exhibit an asymmetric effect for bear and bull markets. Periods with abnormal losses (i.e. either moderate or speculative bear regimes) tend to have a greater variance than those periods in which the NEX increases at faster speed than the explanatory factors (i.e. either moderate or speculative bear regimes).
Inchauspe, State-space risk measurement: An application to renewable energy returns

4. DISCUSSION AND CONCLUSIONS

This paper has contributed to understanding the driving forces of the NEX market and its sources of risks as well as to the empirical methodology of return analysis. Our first-model provides insights about the dynamic influence of different factors but has the disadvantage of presenting a volatility cluster around 2008. Incorporating conditional volatility in a Kalman-filter setup may interfere with the assumptions of the latter, hence a different methodology is proposed for dealing with the weekly dataset. I propose a MSMH-ADL model which has been able to identify four distinctive means and distributions for NEX excess returns to address the following question: what would be the expected returns over a long horizon. This contribution disentangles some questions about the determinants of NEX returns and also provides a tool which could be useful in various ways. First, the MSMH-ADL model can be used for investment analysis over a horizon greater than 140 weeks. Second, if the price of the NEX is identified as undervalued according to the return equations in Table 2, the information about the Markov process (Figures 4 and 5) could help the investor make profitable decisions. Third, after allowing for “animal spirits” driving trends, the results provide some information on how much is explained by “fundamentals” and how large is the variation that cannot be attributed to them. Fourth, these results can be used for portfolio analysis. Many portfolios use commodities such as oil to hedge risk as the oil price typically bears a negative correlation to many assets. Thus, a natural question that arises in portfolio analysis is whether or not the NEX has characteristics that make it similar to oil prices. None of the models considered in this paper have found evidence of excess returns on oil prices being a significant contributor to NEX excess returns. Nevertheless, the second model suggests that both Amex Oil and Nasdaq returns contribute significantly to NEX returns. An intuitive explanation may be that both of these indices list firms that compete for similar inputs. An alternative explanation that has been put forward in the literature says that investors tend to see renewable energy firms as similar to high technology companies (Henriques and Sadorsky, 2008), perhaps lacking enough information about this newly-developed market. For the weekly dataset used in this paper, it is found that the 4-regime Markov switching model yields the best and smoother fit to data with no excessive outliers in its standardized residuals.

ACKNOWLEDGMENTS

I am grateful to the Centre for Research in Minerals and Energy Economics and Curtin University for their support. I acknowledge the contributions by Stefan Trueck and Ronald Ripple to Section 2 of this paper.
REFERENCES


