Unknown unknowns in the German-Nordic 2019 electricity price difference.

Ragnar Nymoen
University of Oslo*

23 November 2018

Abstract

The German-Nordic electricity price difference (future price) is a wide sense non-stationary series, with a unit-root and with intermittent location-shifts. Using appropriate econometric methods, those defining features of the series could have been asserted empirically before the large price difference occurred in September 2018. Hence it is the magnitude of the locations-shifts in the September that makes them different from earlier breaks, not their nature as "unknown unknowns" for any forecaster or investor who attempted to foresee the development of the price difference based on the information in the historical time series. The use of stationary models with ARCH is likely to lead to underestimation of market price volatility, both in the short term and medium term perspective, maybe allowing too large positions being taken by market operators.

1 Introduction

In September 2018, a huge default by an individual trader in the German-Nordic power market led to inquests and demands for fresh funds from market participants, as the Nasdaq exchange and its customers had to rebuild their defences against defaults in the derivatives markets after they were badly damaged by soured bets from an individual trader. Einar Aas, one of Norway's best known power markets traders, was unable to maintain his positions in the German and Nordic power markets after prices went sharply against him. The outsized positions used up several layers of protection at Nasdaq's clearing house, which is designed to insulate the market from the effects of a default. Nasdaq had to use two-thirds of its mutual default fund, in which the market shares any extreme losses. With Nasdaq and members of its clearing house repairing the damage, questions as to how a single trader could come close to wiping out the clearing house's layers of protection attracted the attention of regulators, including the European Central Bank.[1]

Another aspect has to do with the underlying perception of risk and volatility in the market for power contracts, not least from a regulator's point of view. If the distribution function of the German-Nordic price difference is relatively stable, albeit with higher probability of consequential tail observations than the normal distribution, a certain regulatory protocol can be sufficient both for the clearing house and the stability of the wider system. However, if a different model of the series is more realistic, implying much larger inherent volatility and risk than implied by a stationary ARCH model, one can imagine that different regulation is required.

In this note we document evidence supporting the view that the price difference between German and Nordic future price of electricity is a broad sense non-stationary random variable. Two characteristics of broad sense non-stationarity are unit-roots and location shifts.

*Thanks to Terje Erikstad for discussion. The source of the data set analysed on this note is Nasdaq OMX Commodities/Macrobond. The numerical results and graphs is this note were produced by OxMetrics 8.0/PcGive 15.0.Doornik and Hendry (2018)

[1]Financial Times 14 September 2018.

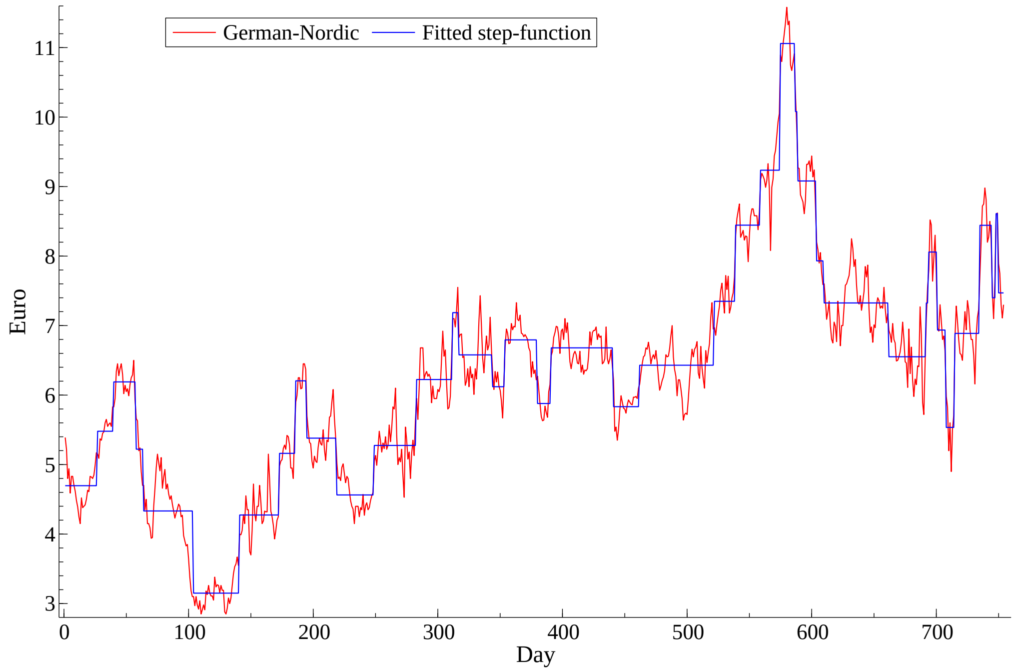

Figure 1: Price difference between German and Nordic future electricity prices, from 28 September 2015 to 30 August 2018.

Already the unit-root property implies that the mean squared forecasts errors converge toward infinity, and that the expected length of a cycle (from peak to through) is as long as the time series itself, making a bet on mean-reversion (closing of the spread) hazardous. The incidence of location shifts in the series add to the unpredictability of the German-Nordic price difference. At any point in time, there a future "unknown unknowns" that are likely to change the price difference in significant ways.

With the aid of a standard econometrics method, and an automatic algorithm for detection of structural breaks, we demonstrate that it would have been possible to identify the presence of both forms of non-stationarity in the German-Nordic price difference, on a sample that ends before the September default. There is also indications that the frequency of location shifts (breaks) became higher though the summer, hence a kind of clustering of risk increasing events would have been detectable during what turned out to be a run-up to the default in the market.

2 Data description

The data set are two time series of daily future prices of electricity, in the German and in the Nordic region. The prices are denoted in EUR and for settlement on 31 December 2019. The two time series start on 28 September 2015, and end on 19 September 2018. The number of observations is 766. Figure 1 shows a time plot of the difference between the German price and the Nordic price from 28 September 2015 to 31 August 2018, the first 754 observations of the price difference, together with a plotted step-function that shows the periods where are relatively stable mean differences between the German and Nordic price.

The graph shows that there are many "mean breaks", or location shifts as we refer to them below, in the time plot. Altogether 38 breaks until 30 August 2018. Hence the average number of consecutive days with a constant expectation of the price difference is estimated to 20 days on the basis of this graph. There are periods where positive and negative breaks

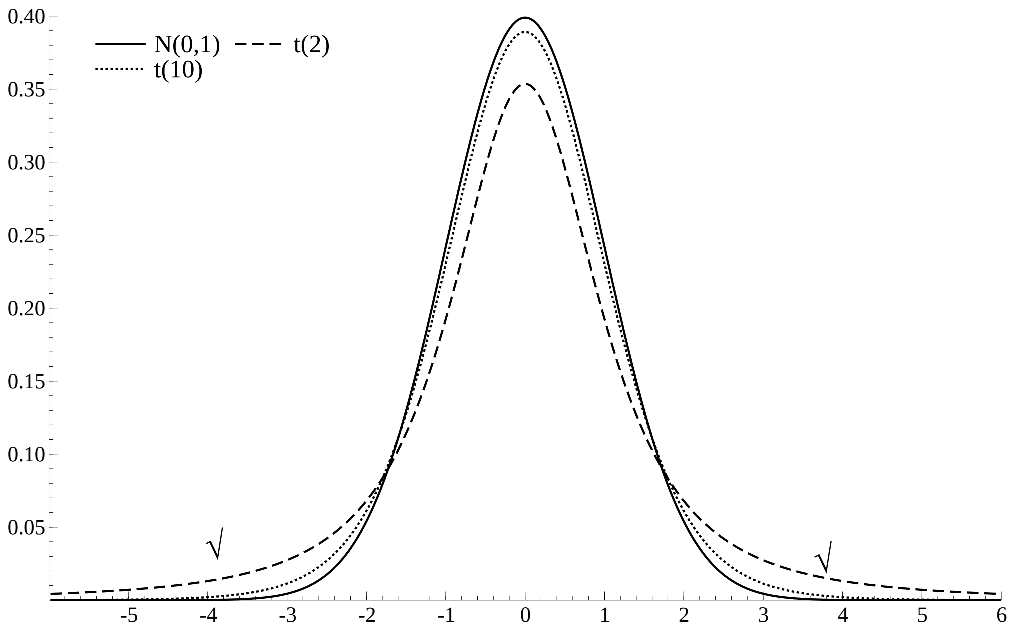

Figure 2: ($N(0,1)$ pdf together with pdf for *Student* t-distribution with 2 and 10 degrees of freedom (with a single black swans indicated in each of the tails)

appear to have been cancelling each other. For example, one can say that a in a longer term view, the average price difference was constant from observation 300 to 500 for example (late November 2016 to the start of September 2017). However, even more striking are the periods where the break are positively autocorrelated, creating the peaked formation in the autumn of 2017 and early winter of 2018 (cf. observation 500-600), and before that, the through around observation 100.

A third observation from the graph is that there is a tendency of the stable-mean periods becoming even shorter at the end of the period shown, in July and August 2018. Hence, when the frequency of change in mean is brought into the picture, there is indication that uncertainty about the location of the price average became large as we approached the late summer of 2108.

3 Method

Formally, each of the pairs of observations consisting of German and Nordic prices are regarded as realizations of bivariate statistical distributions. The whole data set is therefore conceptualized as a realization of a joint probability function for 2×766 random variables. From this starting point we can build models of the joint probability function. In that process, sequential conditioning from the past to the present is an analytical step which simplifies model formulation and estimation a great deal, without any loss of information about the focus relationship between the German and Nordic electricity price.

A crucial step in model building is the assumed form of the joint distribution for the two prices (conditional on the past). In the finance literature, there is a concern that the conventional choice of the Gaussian (normal) distribution tends to under-represent the probability of large events, *ie* the so called tail observations that are far removed to the left or the right of centre (*ie* mean) of the distribution.

Investment strategies or prediction models that assume a normal distribution when the distribution is in fact, heavy-tailed, can lead to financial losses (or gains) because the

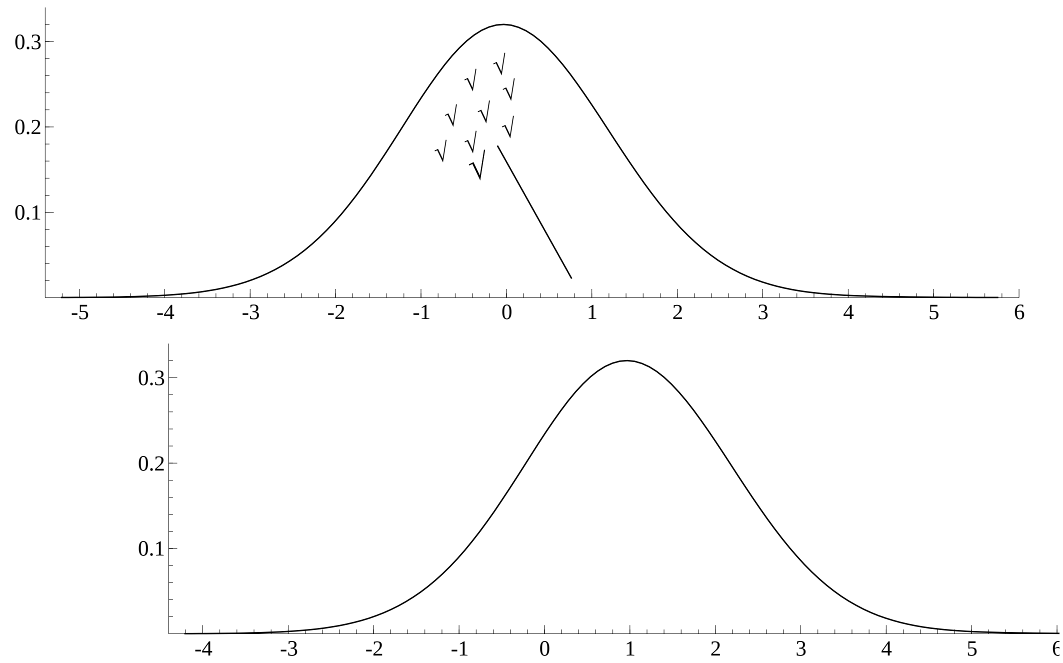

Figure 3: Location-shift (flock of black swans). The upper pdf of $N(0,1)$ gets its location parameter shifted to unity (the pdf in the lower graph). The flight direction of the swans is indicated by the straight line.

probabilities of large impact events are underestimated. It has become known as the *Black Swan Problem*, after the best selling book by Nassim Taleb (2010). Figure 2 gives a simple illustration, with the plotted *Student-t* distribution with 2 degrees of freedom (labelled $t(2)$ in the graph) displaying most clearly the excess probability of rare events.

Of course, models of financial variables that are used in practice do not rely on the constancy of the assumed distribution's parameters. The ARCH model in particular has become a popular one because it can capture features that are recognized as typical of financial variables. Time varying volatility (non constant variance of equity returns) in general, and more specifically clustering of volatility. There are typically periods when large changes in a return index is followed by further large changes and other periods when small changes are followed by further small changes.

The ARCH model (and the specific generalization known as GARCH) explains the volatility clustering of financial returns as a function of the errors of a model that assumes constant mean (the first order moment or location parameter). The observable errors are called "news" or "shocks". However, if the assumed constancy of the location parameter is doubtful, users of these model may be led to attribute too much predictive power the recent history of shocks has about the future development of the returns to stock indices and of other financial variables. A further issue is that the models also have implications for the risk carried by positions in the market, which is a main concern of regulators. The difficult task of deciding on a model of observed volatility features is therefore a consequential one.

Builders and users of macroeconomic models often experience shocks of a different type, namely structural breaks that cause the first moment of a variable's distribution to change, Nymoen (2019, Ch. 11-12). Hence, the black swan analogy needs to be extended to a flock of black swans, that we can imagine lifting the location of a distribution from one point on the real line to another (Hendry (2018), Hendry and Nielsen (2007, p. 32)), as illustrated in Figure 3.

In fact, Figure 1 demonstrated location shifts and can be interpreted as 39 distributions with different means but equal variances. As another illustration of this phenomenon, Figure

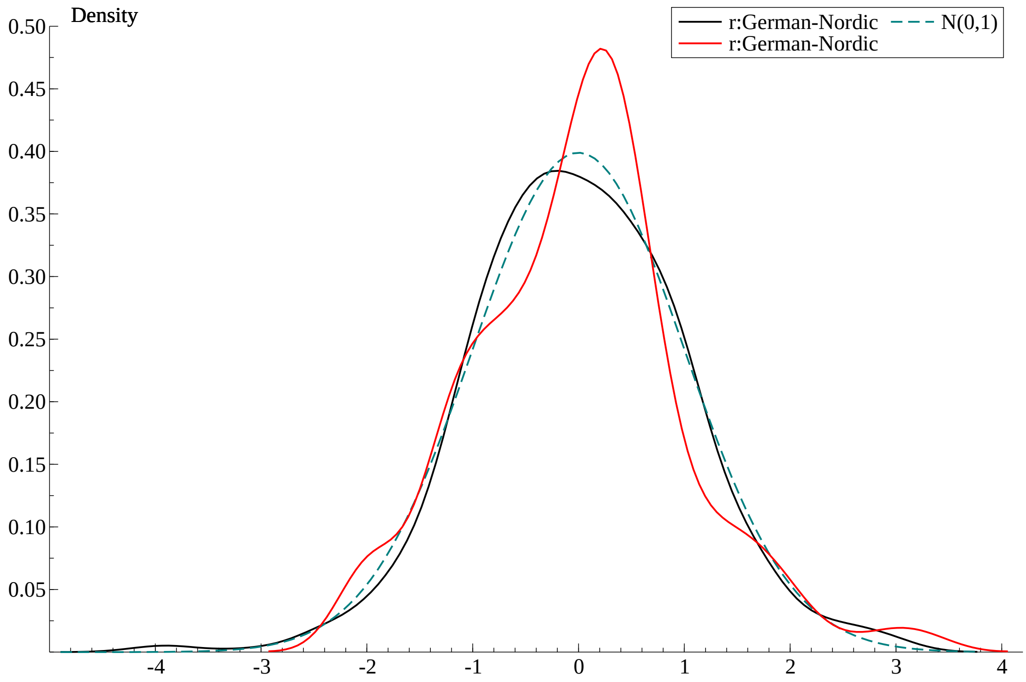

Figure 4: Estimated density function of the German-Nordic price difference (red line) and for the price difference controlled for the location breaks in Figure 1. Data from 28 September 2015 to 30 August 2018

4 shows the estimated probability densities for the raw German-Nordic price difference, and the price difference after controlling for the 38 location breaks in Figure 1. The standard normal density is also plotted for comparison (dashed line). Clearly, the plotted density of the raw price difference shows clear signs of departures from the normal distribution. It appears to be skewed to the right, and there are several "shoulders" consistent with multiple means (imagine putting one of the densities in Figure 3 on top of the other).

The density for the price difference after the removal of multiple location shifts is seen to be much more aligned with the standard normal. There is still a tendency of skewness (but much weaker), and there appears to be more probability in the tails than the normal distribution allows for. Hence there may be a black-swan problem in the price difference series, but it appears to be dominated by the non-stationarity induced by the changes in the mean.

Models that allow for location shifts, and econometric methods that have the power to discover location breaks also seem to have considerable relevance for regulators of some commodity (derivative) markets, as the unlikely huge impact of Mr. Aas' positions on the defense fund of Nasdaq and its members may have demonstrated. The method has been developed in macro econometrics but have found applications in such diverse fields as prediction of volcanic eruption and forecasting of hurricane damage. Finally, the method is complementary to statistical methods that are in use in empirical finance and which focus on the second order moment (the scale parameter) of the time series.

Location shift modelling can be used to identify empirically shocks to markets, or internal behavioural changes, that have changed the expected return of an financial for example, for shorter or longer time periods. Even though it is not risk analysis of the conventional type, again that analysis would focus on the variance of the series, it can inform investors and regulators about changes in the predictability of price developments, which in practice is an important component of risk assessment.

Location shift modelling of the semi-automatized type comes in several versions. The variant dubbed Step Indicator Saturation (SIS) was used create the descriptive step function graph in Figure 1. The version dubbed Impulse Indicator Saturation (IIS) is better suited for modelling the idea that a large price difference today will be become closed wholly of partially by future price changes. This type of dynamic model is well known in econometrics and is called Equilibrium Correction Model, ECM.

We estimate an ECM for the change in the German-Nordic price difference. We use the variable selection algorithm *Autometrics*, which is part of PcGive. To allow for inter-day correlations, we included six lags of the change in the price difference in the list of regressors, in addition to the focus variable which is the lagged price difference (the expected sign of the coefficient is negative). When we estimate this model using IIS, the algorithm starts by estimating a large model, where there is one dummy for each day in the sample. The initial model is saturated by impulse indicators (*ie* dummies).

Since the model then has one indicator variable for each observation in the sample, there are more variables than observations (denote that number by T). However, the algorithm has an elegant solution to the problem of more variables than observations. In the simplest case it is to add the indicators in blocks of $T/2$, noting that all the indicators are mutually uncorrelated. The algorithm then adds half of the indicators to the GUM (*eg* the null model in the simplest case) and selects as usual, records the outcome and drops that first indicator set. Next, add the second set of $T/2$ indicators and select again. Then the retained indicators from the first two selections are combined and added to the GUM, and the selection algorithm is run again as is if we commenced with a number of indicators well below T, see Hendry and Doornik (2014, Ch. 15).

4 Modelling the German-Nordic price difference

Since the drama in the investment market derived from electricity production and distribution occurred in early September 2018, we initially take the September observations out of the sample when we estimate the ECM. In that way we can investigate empirically whether there were flocks of black swans in flight already before the default that shook the market and the clearing house.

When the September prices are dropped from the sample, the number of observations becomes 754. An important decision in automatic variable selection concerns the choice of overall significance level. As a rule of thumb, the number of false positives (concluding with a flock of swans, when in fact there were none) is then $754 \cdot \text{significance level}$. Hence if we set *Significance level*$= 0.001$ the expected number of false positives is less than one (0.754), which seems an acceptable low number.

Using a sample from late September 2015 to the end of August 2018, and calibrated for IIS with *Significance level* $= 0.001$, Autometrics finds 24 indicator variables, as shown in

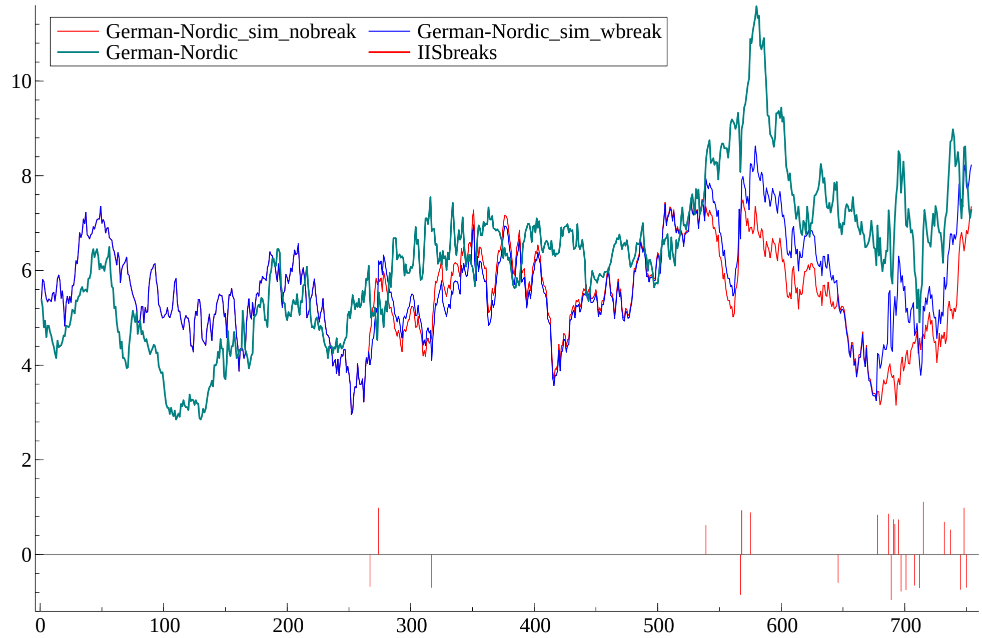

Figure 5: Price difference between German and Nordic future electricity prices, from 28 September 2015 to 30 August 2018, together with simulated series using (4), with and without the use of the indicator variables (shown as bars).

the estimted model equation (4).

$$
\begin{aligned}
\text{German-Nordic} \quad = \quad & \underset{(0.00557)}{0.9869} \text{ German-Nordic}_{t-1} - \underset{(0.234)}{0.6807} \text{ I:}267_t + \underset{(0.235)}{0.9887} \text{ I:}274_t \\
& - \underset{(0.234)}{0.7017} \text{ I:}317_t + \underset{(0.234)}{0.6203} \text{ I:}539_t - \underset{(0.235)}{0.853} \text{ I:}567_t \\
& + \underset{(0.235)}{0.9353} \text{ I:}568_t + \underset{(0.235)}{0.8912} \text{ I:}575_t - \underset{(0.235)}{0.5975} \text{ I:}646_t \\
& + \underset{(0.234)}{0.8394} \text{ I:}678_t + \underset{(0.234)}{0.8634} \text{ I:}687_t - \underset{(0.234)}{0.9607} \text{ I:}689_t \\
& + \underset{(0.234)}{0.7443} \text{ I:}691_t + \underset{(0.234)}{0.6442} \text{ I:}692_t + \underset{(0.235)}{0.7416} \text{ I:}695_t \\
& - \underset{(0.235)}{0.7798} \text{ I:}697_t - \underset{(0.234)}{0.7493} \text{ I:}701_t - \underset{(0.234)}{0.6536} \text{ I:}708_t \\
& - \underset{(0.234)}{0.7073} \text{ I:}712_t + \underset{(0.234)}{1.114} \text{ I:}715_t + \underset{(0.234)}{0.6901} \text{ I:}732_t \\
& + \underset{(0.235)}{0.5271} \text{ I:}737_t - \underset{(0.235)}{0.7426} \text{ I:}745_t + \underset{(0.234)}{0.9894} \; h\text{I:}748_t \\
& - \underset{(0.235)}{0.6976} \text{ I:}750_t + \underset{(0.0355)}{0.08077} \\
& \text{T (No of observations): } 753, \sigma = 0.234168
\end{aligned}
$$

Diagnostic tests support that the error terms of the model equation approximates the Gaussian model. In particular there is no clear evidence of ARCH effects in the errors of the model equation. The test of first order ARCH gets a p-value of 0.02. For higher order ARCH, the p-values are higher than the conventional significance level. Of course, the notable departure from the stationary Gaussian model is that there are 24 estimated changes (breaks) in the mean of the German-Nordic price difference, which conditional on

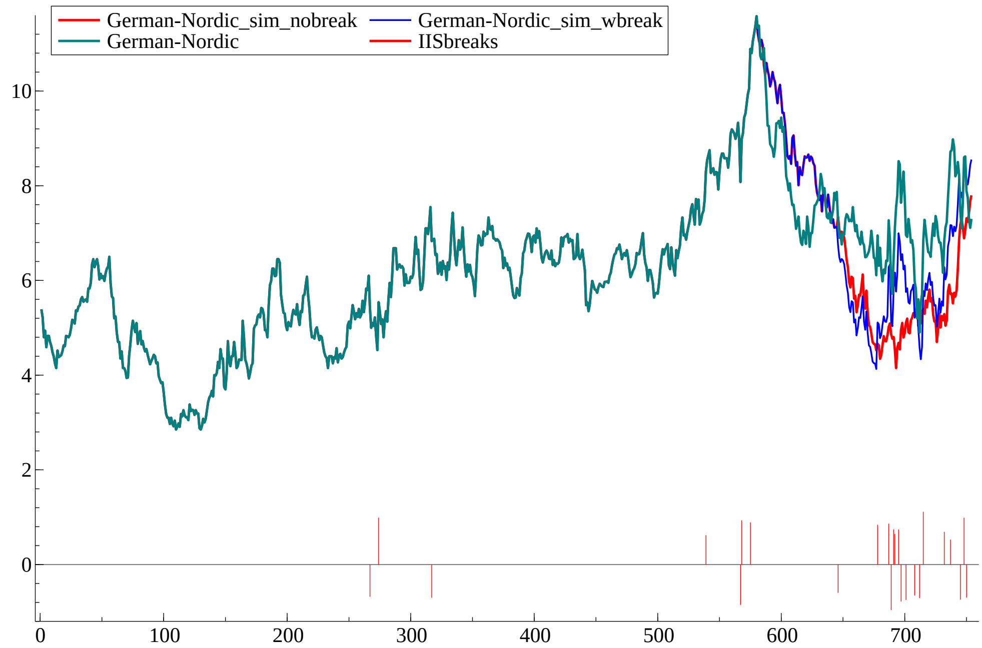

Figure 6: Price difference between German and Nordic future electricity prices, from 28 September 2015 to 30 August 2018, together with simulated series from the peak (27 December 2017) using (4), with and without the use of the indicator variables (shown as bars).

no break is estimated to be:

$$German - Nordic = 6.15298 \text{ Euro} \quad (\text{“on average”}).$$

However, the model estimation results also show that this estimate is not reliable. This is seen from the estimated coefficient of the lagged price difference, which is 0.9869 and insignificantly different from 1 when the correct critical value is used to test the t-value of -2.41. Hence, the correct interpretation is that there is little evidence of mean reversion in the time series for the German-Nordic price difference. The implication is that any gamble on how many days it takes before the price difference “returns to normality” after an increase is a wild speculation, or if not, needs to make makes use of a wider (private) information set, than the information in the data of the price difference itself. Mr. Aas' bet was of exactly this type: He would make money if the price difference became reduced, consistent with the coefficient of $German - Nordic_{t-1}$ being significantly less than one.

Figure 5 shows the actual price difference together with graphs of two simulated series (simulation start on 29 August 2015). The graph labelled German-Nordic_sim_wbreak is for (4) as estimated with the indicators. The graphs labelled German-Nordic_sim_nobreak uses the same equation, but omits the indicators in the simulation. The difference between the two graphs therefore shows the impact of the indicator set, shown as bars in the figure on the solution. Compared to the inherent difficulty of predicting the price differential, the extra “unknown unknowns” do not appear to make much difference. However, this is for a (in this context) very long forecast horizon.

As noted, Figure 5 also shows the break indicators, as bars, and brings out that there is a clustering of indicators for breaks towards the end of the sample period: Of the 24 breaks found between end of September 2015 and end of August 2018, 15 are from June-August 2018.[2] This could be a signal of increasingly volatile series, which a regulator might want

[2] 7 in June, 1 in July and 5 in August.

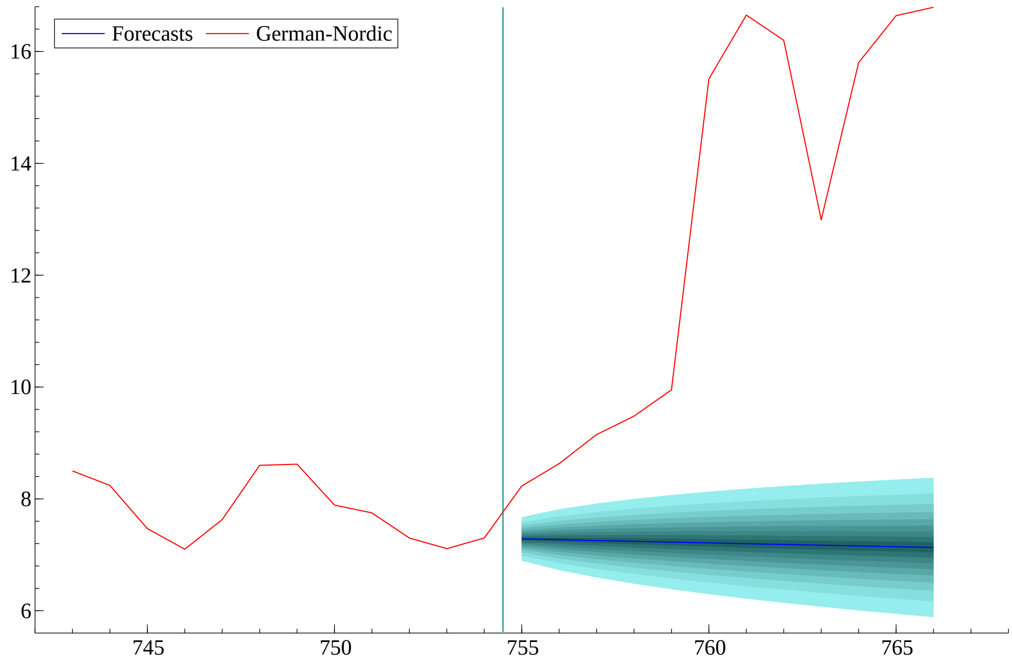

Figure 7: Forecasting German-Nordic price for 12 days in September 2018 based on the last observation from August, using the model equation (4)

to look into. Figure 6 shows the actual and the simulated series when the simulations starts from the peak which occurred on 27 December 2017.

As can be expected, the equilibrium correction nature of the simulated series coincides with the sharp reduction in the price difference over this period, and could maybe encourage a belief in some a degree predictability of the series.

At the end of August, the price difference was a little above the mean of 6.15, meaning that a forecast from that initial condition would be a slight reduction in the price difference. However, the German price then increased and Norwegian future prices fell as an unusually dry summer came to an end. Figure 7 shows the forecasted price difference and a fan-chart (95 % forecast interval). Clearly, this is a massive forecast failure, which is due in part to the starting date which happens to move the forecast down as noted, but which in the main is due to the underestimation of the real forecasting uncertainty for the German-Nordic price difference.

Re-estimating of the model equation, Autometrics with IIS finds 8 breaks in 12 day period at the start of September, some of very large as already evident from the graphs.

$$
\begin{aligned}
\text{German-Nordic} \quad = \quad & \underset{(0.00512)}{0.9937} + \text{earlier breaks} \\
& + \underset{(0.223)}{0.5317}\ \text{I:}757_t \ + \ \underset{(0.223)}{0.487}\ \text{I:}759_t \\
& + \ \underset{(0.224)}{5.58}\ \text{I:}760_t \ + \ \underset{(0.228)}{1.195}\ \text{I:}761_t \ - \ \underset{(0.229)}{3.15}\ \text{I:}763_t \\
& + \ \underset{(0.225)}{2.849}\ \text{I:}764_t \ + \ \underset{(0.228)}{0.897}\ \text{I:}765_t \ + \ \underset{(0.0328)}{0.04295} \\
& \text{T (No of observations): } 766, \sigma = 0.223
\end{aligned}
$$

5 Conclusions

A relevant model of the German-Nordic price difference of electricity (forward-price) appears to be a wide sense non-stationary model, with a unit-root and with intermittent location-shifts (flocks of black swans). The popular ARCH model seems to entail too much stationarity to be a realistic model for the price difference time series. Using appropriate econometric methods, these defining features of the series could have been asserted empirically, before the large price increases in early September. Hence, it is the magnitudes of the locations-shifts in the September that make them different from earlier breaks, not their nature as "unknown unknowns" for any forecaster or investor who attempted to foresee the development of the price difference based on the information in the historical time series. Individual operators in the markets for derivatives use private information when they decide their forecasts. However, regulators need to use statistical models to characterize the risk carried by existing and hypothetical positions in the market. Basing that assessment on stationary models with ARCH effects is likely to lead to underestimation of volatility, and may allow too large positions to be taken by market operators.

References

Doornik, J. A. and D. F. Hendry (2018). *Modelling Dynamic Systems PcGive 15. Volume 1*. Timberlake Consultants, London.

Hendry, D. F. (2018). Deciding Between Alternative Approaches in Macroeconomics. *International Journal of Forecasting*, *34*(1), 119–135.

Hendry, D. F. and J. A. Doornik (2014). *Empirical Model Discovery and Theory Evaluation. Automatic Selection Methods in Econometrics*. Arne Ryde Memorial Lectures. MIT Press, Cambridge,Mass.

Hendry, D. F. and B. Nielsen (2007). *Econometric Modeling*. Princeton University Press, Princeton.

Nymoen, R. (2019). *Dynamic Econometrics for Empirical Macroeconomic Modelling*. World Scientific, Boston.

Taleb, N. N. (2010). *The Black Swan: The Impact of the Highly Improbable: With a new section: "On Robustness and Fragility"*. Penguiin Random House, London, 2nd edn.