FOLLOW-UP NOTE ON MARKET STATE MODELS

In an earlier note I outlined some of the available techniques used for modeling market states. The following is an illustration of how these techniques can be applied in practice.

The chart below shows the daily compounded returns for a single pair in an ETF statistical arbitrage strategy, back-tested over a 1-year period from April 2010 to March 2011.

The idea is to examine the characteristics of the returns process and assess its predictability.

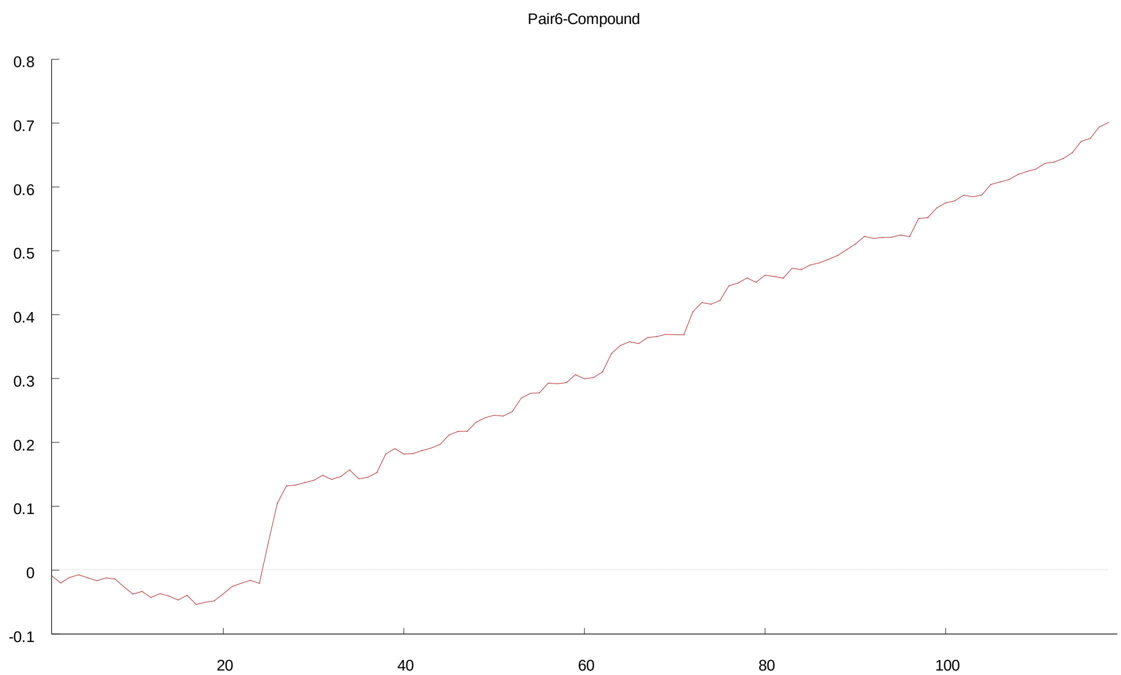

FIG 1. Compound Daily Returns in Pair6

The initial impression given by the analytics plots of daily returns, shown in Fig 2 below, is that the process may be somewhat predictable, given what appears to be a significant 1-order lag in the autocorrelation spectrum. We also see evidence of the customary non-Gaussian "fat-tailed" distribution in the error process.

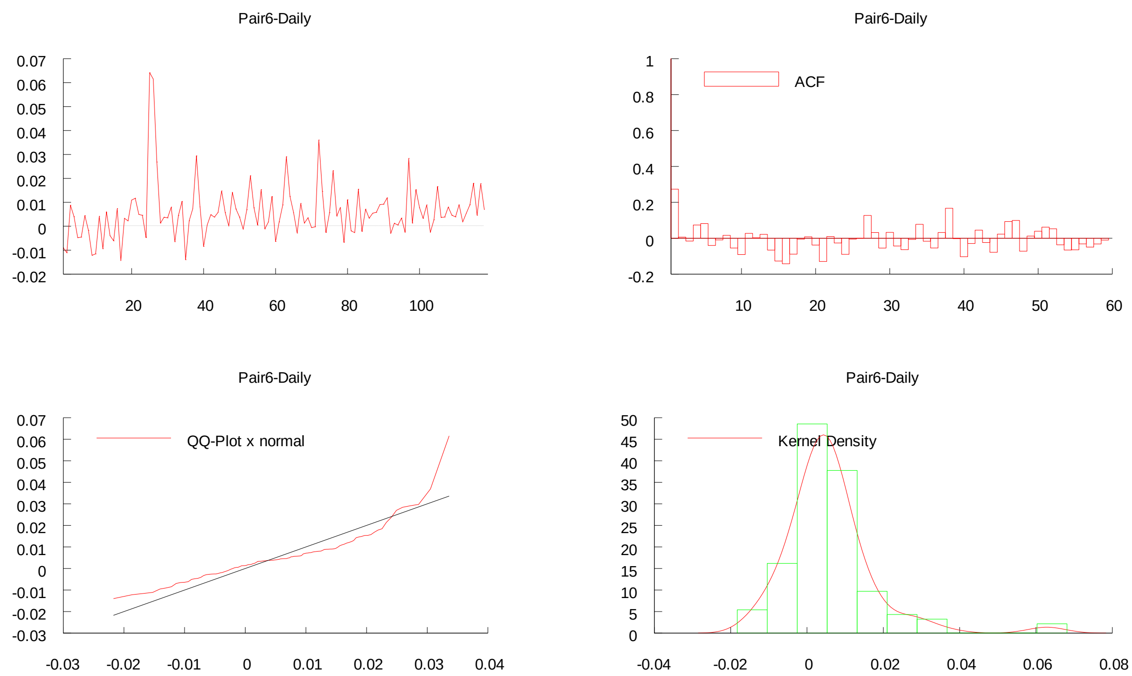

FIG 2. Analytical Plots of Daily Returns for Pair6

An initial attempt to fit a standard Auto-Regressive Moving Average ARMA(1,0,1) model yields disappointing results, with an unadjusted model R-squared of only 7% (see model output in Appendix 1)

However, by fitting a 2-state Markov model we are able to explain as much as 65% in the variation in the returns process (see Appendix II).

The model estimates Markov Transition Probabilities as follows.

	P(.|1)	P(.|2)
P(1|.)	0.93920	0.69781
P(2|.)	0.060802	0.30219

In other words, the process spends most of the time in State 1, switching to State 2 around once a month, as illustrated in Fig 3 below.

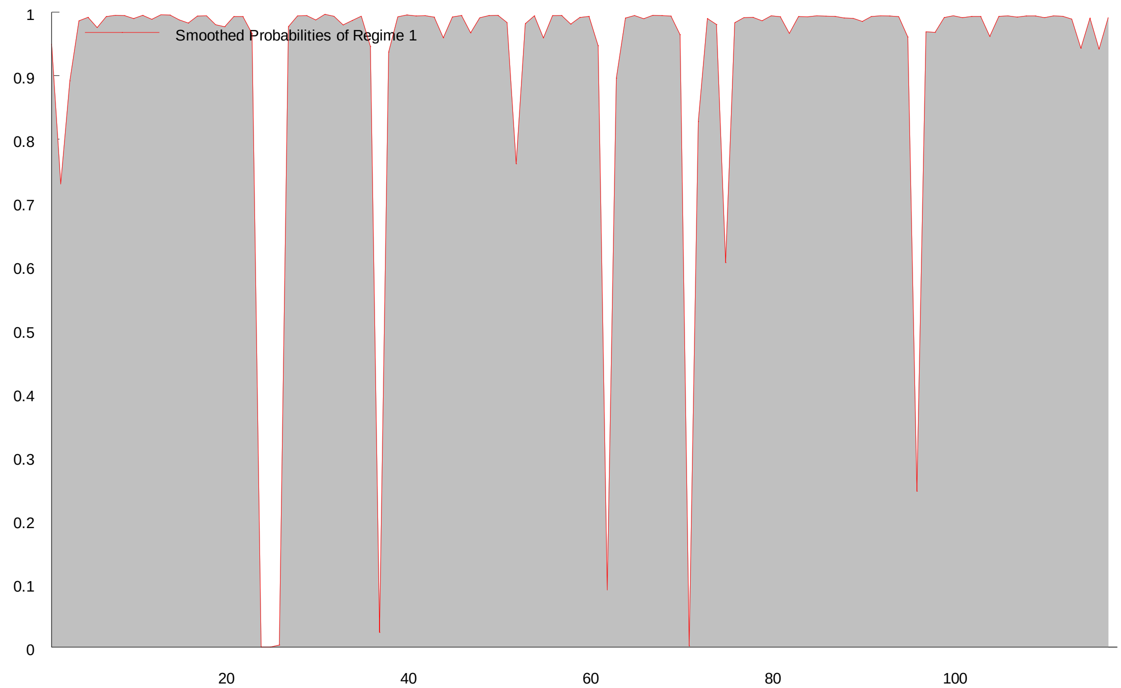

FIG 3. Smooth State Transition Probabilities

In the first state, the pairs model produces an expected daily return of around 65bp, with a standard deviation of similar magnitude. In this state, the process also exhibits very significant auto-regressive and moving average features.

Regime 1:

Intercept	0.00648	0.0009	7.2	0
AR1	0.92569	0.01897	48.797	0
MA1	-0.96264	0.02111	-45.601	0
Error Variance^(1/2)	0.00666	0.0007	------	------

In the second state, the pairs model produces lower average returns, and with much greater variability, while the autoregressive and moving average terms are poorly determined.

Regime 2:

Intercept	0.03554	0.04778	0.744	0.459
AR1	0.79349	0.06418	12.364	0
MA1	-0.76904	0.51601	-1.49	0.139

Error Variance^(1/2)   0.01819   0.0031   ------   ------

CONCLUSION

The analysis in Appendix II suggests that the residual process is stable and Gaussian. In other words, the two-state Markov model is able to account for the non-Normality of the returns process and extract the salient autoregressive and moving average features in a way that makes economic sense.

How is this information useful? Potentially in two ways:

(i) If the market state can be forecast successfully, we can use that information to increase our capital allocation during periods when the process is predicted to be in State 1, and reduce the allocation at times when it is in State 2.

(ii) By examining the timing of the Markov states and considering different features of the market during the contrasting periods, we might be able to identify additional explanatory factors that would be used to further enhance the trading model.

FIG 4. In-Sample and Out-Of-Sample Forecasts

APPENDIX I - ARMA(1,0,1) MODEL

Dependent Variable is Pair6-Daily

117 observations (2-118) used for estimation

with 1 pre-sample observations

and 61 ex-post forecasts.

Estimation Method: Conditional ML (Time Domain)

Gaussian Likelihood

ARIMA(1,0,1)

Strong convergence

iteration time: 0.04

	Estimate	Std. Err.	t Ratio	p-Value
Intercept	0.00609	0.00142	4.29	0
AR1	0.2172	0.20388	1.065	0.289
MA1	0.06548	0.32787	0.2	0.842
Error Variance^(1/2)	0.01109	0.0017	------	------

Log Likelihood = 360.665

Schwarz Criterion = 351.141

Hannan-Quinn Criterion = 354.422

Akaike Criterion = 356.665

Sum of Squares = 0.0144

R-Squared = 0.0763

R-Bar-Squared = 0.0517

Residual SD = 0.0105

Residual Skewness = 1.995

Residual Kurtosis = 10.8968

Jarque-Bera Test = 381.611 {0}

Box-Pierce (residuals): Q(23) = 10.9463 {0.984}

Box-Pierce (squared residuals): Q(25) = 13.1709 {0.974}

Forecast Test 1: ChiSq(61) = 36.5236 {0.995}

Forecast Test 2: N(0,1) = -1.2098 {0.226}

Listing saved in Series94.xls

MA form is 1 + a_1 L +...+ a_q L^q.

Covariance matrix from robust formula.

APPENDIX II - 2-STATE MARKOV ARMA(1,0,1) MODEL

117 observations (2-118) used for estimation

with 1 pre-sample observations

and 61 ex-post forecasts.

Estimation Method: Conditional ML (Time Domain)

Gaussian Likelihood

Markov-switching model with 2 regimes

ARIMA(1,0,1)

Strong convergence

iteration time: 1.07

Markov Transition Probabilities

	P(.|1)	P(.|2)
P(1|.)	0.93920	0.69781
P(2|.)	0.060802	0.30219

	Estimate	Std. Err.	t Ratio	p-Value
Logistic, t(1,1)	2.73741	0.8215	------	------
Logistic, t(1,2)	0.8369	0.5309	------	------

Non-switching parameters shown as Regime 1.

Regime 1:

Intercept	0.00648	0.0009	7.2	0

AR1	0.92569	0.01897	48.797	0
MA1	-0.96264	0.02111	-45.601	0
Error Variance^(1/2)	0.00666	0.0007	------	------

Regime 2:

Intercept	0.03554	0.04778	0.744	0.459
AR1	0.79349	0.06418	12.364	0
MA1	-0.76904	0.51601	-1.49	0.139
Error Variance^(1/2)	0.01819	0.0031	------	------

Log Likelihood = 391.357

Schwarz Criterion = 367.546

Hannan-Quinn Criterion = 375.75

Akaike Criterion = 381.357

Sum of Squares = 0.0059

R-Squared = 0.653

R-Bar-Squared = 0.6238

Residual SD = 0.0081

Residual Skewness = -0.2063

Residual Kurtosis = 2.5963

Jarque-Bera Test = 1.6245 {0.444}

Box-Pierce (residuals): Q(23) = 21.4747 {0.552}

Box-Pierce (squared residuals): Q(25) = 21.172 {0.683}

Forecast Test 1: ChiSq(61) = 76.6542 {0.085}

Forecast Test 2: N(0,1) = 0.9587 {0.338}

Listing saved in Series92.xls

MA form is 1 + a_1 L +...+ a_q L^q.

Covariance matrix from robust formula.