# Markov Regime-Switching (and some State Space) Models in Energy Markets

Matthew Brigida, Ph.D.

Department of Finance
College of Business Administration
Clarion University of Pennsylvania
Clarion, PA 16214
mbrigida@clarion.edu

June 2015

# Introduction

Time series often exhibit distinct changes in regime. Thus we must allow for switches in model parameters and standard errors. For example events which may cause a distinct shift are:

- recession or economic expansion (any economic variable or relationship).
- central bank currency intervention (currency volatility).
- the introduction of hydrofracking (the relationship between natural gas and crude oil prices).
- a weather event or plant outage (power price volatility, average returns, autocorrelation).

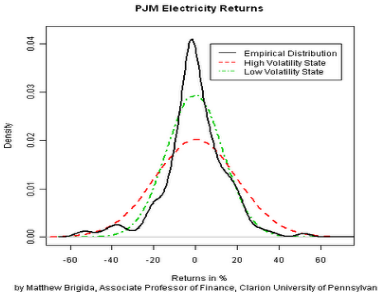
However, we usually don't know exactly when the event occurred, so we cannot use introduce indicator variables. Moreover, it is preferable to have a probability law over the entire data generating process (because *a priori* the timing of the regime switch is uncertain).

# Introduction

Regime-switching may also explain deviations from normality often seen in time series.

- Do stock returns really have fat tails (motivating a Cauchy type distribution)? Or rather are returns normal, but generated by multiple regimes?
- Skewness may be explained similarly.

# An Example: Electricity Prices




$r_t = \mu_{S_t} + e_{S_t}, \quad e_{S_t} \sim N(0, \sigma^2_{S_t})$

Low volatility regime: $\mu_L$ = -0.34%, $\sigma_L$ = 13.5%

High volatility regime: $\mu_H$ = 0.2%, $\sigma_H$ = 19.8%

Empirical: $\mu$ = -0.12%, $\sigma$ = 16.3%, skewness = 0.354, kurtosis = 4.8

The transition matrix for the Markov process is:

| | Low | High |
|---|---|---|
| Low | 63% | 34% |
| High | 58% | 41% |

# Introduction

Ultimately, we will allow the time series to follow $n$ distinct density functions, and switch between these density functions according to an unobserved Markov process. Thus, we need to make inference about the parameters of the density functions, as well as the Markov process (transition probabilities and filtered (or smoothed) regime probabilities $P(S_t = i|\varphi_{t-1})$)[1].

[1] where $\varphi_{t-1}$ is the information available through time $t-1$.

# Gregory and Hansen (1996) Test for Regime-Shifts in a Cointegrating Relationship

Prior to using a Markov-switching model it is useful to test for regime switching.

### Code

```
oil <- log(data$oil)
gas <- log(data$ng)

ind <- matrix(0, nrow=length(oil), ncol=length(oil))

for(i in 1:length(oil)){
  ind[,i] <- c(rep(0,i),rep(1,(length(oil)-i)))}

adf.m2 <- 0
for(i in 1:length(oil)){
  adf.m2[i] <- adf.test(lm(gas~ind[,i]+oil)$resid)$statistic}

min(adf.m2)
```

You then compare the minimum ADF statistic to approximate asymptotic critical values supplied in Gregory and Hansen (1996). The critical values were found via simulation.

# Background

Say we have $y_t = \mathbf{x}_t\boldsymbol{\beta} + e_t$, where $e_t \sim i.i.d.N(0, \sigma^2)$, and $\mathbf{x}_t$ is a $k$ vector of exogenous variables. Let $\boldsymbol{\theta}$ denote the vector of parameters (here $\boldsymbol{\theta} = (\boldsymbol{\beta}\ \sigma^2)'$).

- In the non-switching case we may simply maximize the log-likelihood $\mathcal{L}(\boldsymbol{\theta}) = \sum_{t=1}^{T} ln(f(y_t))$ where $f(\bullet)$ denotes the normal distribution.
- If there is switching, and the regime at each time $(S_t)$ is known *a priori* then we have $\mathcal{L}(\boldsymbol{\theta}) = \sum_{t=1}^{T} ln(f(y_t|S_t))$ where the $S_t$ are indicator variables in $\boldsymbol{\beta}_{S_t} = \boldsymbol{\beta}_1(1 - S_t) + \boldsymbol{\beta}_2 S_t$ and $\sigma^2_{S_t} = \sigma^2_1(1 - S_t) + \sigma^2_2 S_t$ where $S_t \in \{0, 1\}$ (in the case of 2 regimes).

# Background: Unobserved Regimes

If the regime states are unobserved, however, then we'll need the joint density $f(y_t, S_t|\varphi_{t-1}) = f(y_t|S_t, y_t)f(S_t|\varphi_{t-1})$ where $\varphi_{t-1}$ denotes the information set though time $t-1$.

- We may then obtain the marginal $f(y_t|\varphi_{t-1})$ by summing over all possible regimes. In the case of two regimes:
  $f(y_t|\varphi_{t-1}) = \sum_{S_t=1}^{2} f(y_t|S_t, \varphi_t)f(S_t|\varphi_{t-1})$
- We then have $\mathcal{L}(\boldsymbol{\theta}) = \sum_{t=1}^{T} ln(f(y_t|\varphi_{t-1}))$.
  - We then maximize this wrt $\boldsymbol{\theta} = (\boldsymbol{\beta}_{S_1}\ \boldsymbol{\beta}_{S_2}\ \sigma^2_{S_1}\ \sigma^2_{S_2})'$ and **whatever** parameters govern the regime switching process.

# The Regime Switching Process

Now we must consider the process governing regime-switching (i.e. $f(S_t|\varphi_{t-1})$). We may consider:

- Switching which is independent of prior regimes (can be dependent on exogenous variables).
- Markov-switching with constant transition probabilities (dependent on the prior or lagged regime).
- Markov-switching with time-varying transition probabilities (the regime is a function of other variables[2]).

[2] the variables must be conditionally uncorrelated with the regime of the Markov process (Filardo (1998))

# Independent Switching

In the case of independent switching $f(S_t|\varphi_{t-1})$ may have the law:

- $P[S_t = 1] = \frac{e^p}{1+e^p}$
- $P[S_t = 2] = 1 - \frac{e^p}{1+e^p}$

where $\mathcal{L}$ is maximized (unconstrained) wrt $\boldsymbol{\theta} = (\boldsymbol{\beta}_{S_1}\ \boldsymbol{\beta}_{S_2}\ \sigma^2_{S_1}\ \sigma^2_{S_2}\ p)'$. That is the parameters governing $f(S_t|\varphi_{t-1})$ which is $p$, and the parameters governing $f(y_t|S_t, \varphi_t)$ which are $\boldsymbol{\beta}_{S_1}$ , $\boldsymbol{\beta}_{S_2}$ , $\sigma^2_{S_1}$ , $\sigma^2_{S_2}$

You could also replace $p$ with a function of exogenous variables $(z_{t-1})$:

- say, replace $p$ with $p_0 + p_1 z_{1,t-1} + p_2 z_{2,t-1}$ and maximize wrt $p_0$, $p_1$, $p_2$.

This formulation is often implausible from an economic standpoint. Usually the present regime is dependent on the prior regime (if we are in a recession today, there is a greater probability of a recession tomorrow).

# Markov Regime-Switching: Constant Transition Probabilities

Continuing with 2 regimes, we'll allow the regimes to evolve according to a first-order Markov process with constant transition probabilities:

$$\mathbf{P} = \begin{bmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{bmatrix}.$$

- We'll need to make an optimal forecast, and optimal inference, of the regime at each $t$ (this is where the Hamilton filter comes in).
  - That is, we'll need $P[S_t = j|\varphi_{t-1}, \boldsymbol{\theta}]$ and $P[S_t = j|\varphi_t, \boldsymbol{\theta}]$ where $j \in \{1, 2\}$.

For the inference note:

$$P(S_t = j|\varphi_t; \boldsymbol{\theta}) = P(S_t = j|y_t, \mathbf{x}_t, \varphi_{t-1}; \boldsymbol{\theta}) = \frac{P(y_t, S_t = j|\mathbf{x}_t, \varphi_{t-1}; \boldsymbol{\theta})}{f(y_t|\mathbf{x}_t, \varphi_{t-1}; \boldsymbol{\theta})}$$

$$= \frac{P(S_t = j|\mathbf{x}_t, \varphi_{t-1}; \boldsymbol{\theta}) f(y_t|S_t = j, \mathbf{x}_t, \varphi_{t-1}; \boldsymbol{\theta})}{\sum_{j=1}^{2} P(S_t = j|\mathbf{x}_t, \varphi_{t-1}; \boldsymbol{\theta}) f(y_t|S_t = j, \mathbf{x}_t, \varphi_{t-1}; \boldsymbol{\theta})}$$

# Continued...

Collect $P(S_t = j|\varphi_t; \boldsymbol{\theta})$ for $j \in \{1, 2\}$ into a [2x1] vector $\hat{\boldsymbol{\xi}}_{t|t}$. Similarly let $\boldsymbol{\eta}_t = f(y_t|S_t = j, \mathbf{x}_t, \varphi_{t-1}; \boldsymbol{\theta})$ be a [2x1] vector. Then we may rewrite the optimal inference in vector notation as (using Hamilton's notation): $\hat{\boldsymbol{\xi}}_{t|t} = \frac{\hat{\boldsymbol{\xi}}_{t|t-1} \odot \boldsymbol{\eta}_t}{(1\ 1)(\hat{\boldsymbol{\xi}}_{t|t-1} \odot \boldsymbol{\eta}_t)}$ where $\odot$ denotes element-by-element multiplication.

- Then

$$\hat{\boldsymbol{\xi}}_{t+1|t} = \begin{bmatrix} P(S_{t+1} = 1|\varphi_t; \boldsymbol{\theta}) \\ P(S_{t+1} = 2|\varphi_t; \boldsymbol{\theta}) \end{bmatrix} =$$

$$= \begin{bmatrix} p_{11} P(S_t = 1|\varphi_t; \boldsymbol{\theta}) + p_{21} P(S_t = 2|\varphi_t; \boldsymbol{\theta}) \\ p_{12} P(S_t = 1|\varphi_t; \boldsymbol{\theta}) + p_{22} P(S_t = 2|\varphi_t; \boldsymbol{\theta}) \end{bmatrix} =$$

$$= \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} P(S_t = 1|\varphi_t; \boldsymbol{\theta}) \\ P(S_t = 2|\varphi_t; \boldsymbol{\theta}) \end{bmatrix} = \mathbf{P}\hat{\boldsymbol{\xi}}_{t|t}$$

  and so $\hat{\boldsymbol{\xi}}_{t+1|t} = \mathbf{P}\hat{\boldsymbol{\xi}}_{t|t}$ is the optimal forecast.

From here we use $\hat{\boldsymbol{\xi}}_{t+1|t}$ to get $\hat{\boldsymbol{\xi}}_{t+1|t+1}$ and the filter proceeds.

# The Log Likelihood

The above filter provides the log likelihood by:

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{t=1}^{T} log((1\ 1)(\hat{\boldsymbol{\xi}}_{t|t-1} \odot \boldsymbol{\eta}_t)) = \sum_{t=1}^{T} log(f(y_t|\mathbf{x}_t, \varphi_{t-1}; \boldsymbol{\theta}))$$

which can be maximized wrt $\boldsymbol{\theta} = (\boldsymbol{\beta}_{S_1}\ \boldsymbol{\beta}_{S_2}\ \sigma^2_{S_1}\ \sigma^2_{S_2}\ p_{11}\ p_{22})'$.

# Estimating Regime Probability

Once you have the estimated parameters ($\boldsymbol{\theta}$) you may use the data and parameters to estimate the probability of each regime at each $t$. You use the same filtering algorithm.

# Flowchart of the Markov Regime Switching Model

Run through Hamilton Filter

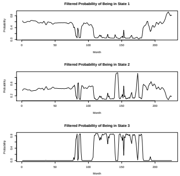
From this we get the Likelihood

$\underset{\boldsymbol{\theta}}{max} \mathcal{L}(\boldsymbol{\theta})$

$\hat{\boldsymbol{\theta}}$ and Standard Errors

Run through Hamilton Filter again with $\hat{\boldsymbol{\theta}}$

Regime Prob.
$P[S_t = j|\varphi_t, \hat{\boldsymbol{\theta}}]$



# Markov-Switching with Constant Transition Probabilities: An Example

There is are theoretical justification for cointegration between natural gas and crude oil prices.

- They are produced together and often sold in contracts which price NG as a fraction of crude oil.
- They are possible substitutes.

However, prior analyses have found conflicting results.

- Some find evidence for cointegration, while others find structural breaks in the relative pricing relationship (prompting claims of 'decoupling').

My paper[3] claims there are events (technological such as fracking and CCCT, and political such as NG deregulation) which cause distinct structural breaks in the cointegrating relationship.

- These breaks may be modeled as switches between regimes.

[3] revise and resubmit at *Energy Economics*

# Markov-Switching with Constant Transition Probabilities: An Example

NG and crude oil cointegrating equation: First-Order, M-Regime, Markov-Switching

$$P_{HH} = \beta_{0,S_t} + \beta_{1,S_t} P_{WTI} + e_t, \quad e_t \sim N\left(0, \sigma^2_{S_t}\right)$$

$$P(S_t = j | S_{t-1} = i) = p_{ij},\ \forall j \in 1, 2, ..., M,\ \text{and} \sum_{j=1}^{M} p_{ij} = 1$$

$$\beta_{0,S_t} = \beta_{0,1} S_{1t} + \beta_{0,2} S_{2t} + ... + \beta_{0,M} S_{Mt}$$

$$\beta_{1,S_t} = \beta_{1,1} S_{1t} + \beta_{1,2} S_{2t} + ... + \beta_{1,M} S_{Mt}$$

$$\sigma_{0,S_t} = \sigma_{0,1} S_{1t} + \sigma_{0,2} S_{2t} + ... + \sigma_{0,M} S_{Mt}$$

$$\text{where for } m \in 1, 2, ...M,\ \text{if } S_t = m \Rightarrow S_{mt} = 1,\ \text{and } S_{mt} = 0 \text{ else}$$

# Estimation

- Construction of the likelihood function for the above Markov Switching cointegrating equation was done using the Hamilton filter[4].
- Minimization of the negative log-likelihood was done using the *optim* function in the R.
- Alternativelly, maximization can be done using the EM algorithm.
  - Closed-form solutions for parameters means no optimization required in M-step.
  - E-step is simply the smoothed probabilities of the unobserved regime, $P(S_t = j|\varphi_T)$.
  - This method is robust with respect to the initial starting values of the model parameters.

[4] see Hamilton (1998).

# R Code

## Log-Likelihood Function

```
lik <- function(theta,
lnoil, lnng){

alpha1 <- theta[1]
alpha2 <- theta[2]
alpha3 <- theta[3]
alpha4 <- theta[4]
alpha5 <- theta[5]
alpha6 <- theta[6]
p11 <- 1 / (1 + exp(-theta[7]))
p22 <- 1 / (1 + exp(-theta[8]))

dist.1 <- 0
dist.1 <- (1/(alpha5*sqrt(2*pi)))*exp((-(lnng-alpha1-alpha3*lnoil)^2)/(2*alpha5^2))

dist.2 <- 0
dist.2 <- (1/(alpha6*sqrt(2*pi)))*exp((-(lnng-alpha2-alpha4*lnoil)^2)/(2*alpha6^2))

dist <- cbind(dist.1, dist.2)
o.v <- c(1,1)
P <- matrix(c(p11, 1-p11, 1- p22, p22), nrow=2, ncol=2)

xi.a <- rep(0,2*length(lnoil))
xi.a <- matrix(xi.a, nrow=(length(lnoil)),ncol=2)
xi.b <- rep(0,2*length(lnoil))
xi.b <- matrix(xi.a, nrow=(length(lnoil)),ncol=2)
model.lik <- rep(0, length(lnoil))
```

# R Code

## Log-Likelihood Function: Continued

```
xi.a[1,] <- (c(p11,p22)*dist[1,])/(o.v%*%(c(p11,p22)*dist[1,]))

## Here is the Hamilton filter

for (i in 1:(length(lnoil)-1)){
  xi.b[i+1,] <- P%*%xi.a[i,]
  xi.a[i+1,] <- (xi.b[i+1,]*dist[i+1,])/(o.v%*%(xi.b[i+1,]*dist[i+1,]))
  model.lik[i+1] <- o.v%*%(xi.b[i+1,]*dist[i+1,])
}

logl <- sum(log(model.lik[2:length(model.lik)]))
return(-logl)
}
```

## Minimization: *optim*

```
theta.start <- c(-.05, .01, .2, .4, .1, .2, .5, .5)
max.lik.optim <- optim(theta.start, lik, lnoil=lnoil, lnng=lnng, hessian=T)

OI <- solve(max.lik.optim$hessian)
se <- sqrt(diag(OI))
t <- max.lik.optim$par/se
pval <- 2*(1-pt(abs(t), nrow(data)-2))
```

# Getting regime probabilities

Once we have found maximum likelihood estimates of the parameters, we rerun the filter with these estimates to get filtered regime probabilities.

## Estimating Regime Probabilities

```
alpha.hat <- c(max.lik.optim$par[1], max.lik.optim$par[2], max.lik.optim$par[3], max.lik.optim$par[4],
max.lik.optim$par[5], max.lik.optim$par[6])

p.0.hat <- c(( 1 / (1 + exp(-max.lik.optim$par[7]))), ( 1 / (1 + exp(-max.lik.optim$par[8]))))
## [...more code similar to above...]
for (i in 1:(length(lnoil)-1)){
  xi.b.hat[i+1,] <- P.hat%*%xi.a.hat[i,]
  xi.a.hat[i+1,] <- (xi.b.hat[i+1,]*dist.hat[i+1,])/(o.v%*%(xi.b.hat[i+1,]*dist.hat[i+1,]))
  model.lik.hat[i+1] <- o.v%*%(xi.b.hat[i+1,]*dist.hat[i+1,])
}
plot(xi.a.hat[,2], type='l', xlab='Month', ylab='Probability',
main='Filtered Probability of Being in Regime 2')
```

# Online app and code

Shinyapp: `https://mattbrigida.shinyapps.io/MSCP`

Code for app (GitHub): ‘Matt-Brigida/R-Finance-2015-MSCP’

# Regime-Weighted Residuals

- Given estimated model parameters and transition probabilities, I calculate the *filtered*[5] regime probability $P(S_t = i|\varphi_{t-1})$ where $\varphi_{t-1}$ denotes the information set at time $t-1$.

- Note this also affords the expected duration of each regime:

$$E(D_m) = \sum_{j=1}^{\infty} j(P[D_m = j]) = \frac{1}{1-p_{mm}}$$

[5]as opposed to smoothed probabilities $P(S_t = i|\varphi_T)$

# Results: 2 regimes & monthly prices

**Table 1. Results of Estimating Cointegrating Equation Regressions (equations 1-5) with Monthly Prices**

There are 225 monthly observations for natural gas and oil prices. The one-state model has 222 degrees-of-freedom, and the two-state model has 217. The alternative hypothesis in the augmented Dickey-Fuller and Phillips-Perron tests is to conclude the residual series is stationary. Residuals in the two-state model are weighted by the filtered state probability. p-values are below the coefficient in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

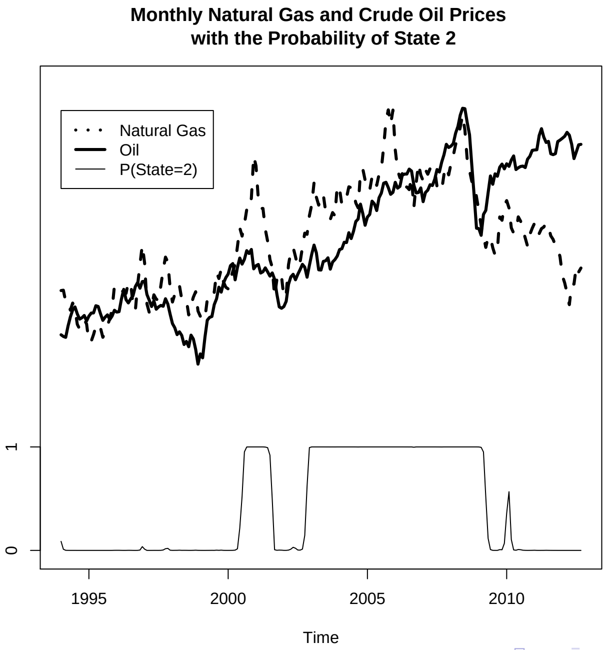
| | **One-State Model** | **Two-State Model** | |
|---|---|---|---|
| State | 1 | 1 | 2 |
| $\beta_0$ | -0.6910<br>(5e-6)*** | -0.0287<br>(0.7518) | 0.3351<br>(0.2068) |
| $\beta_1$ | 0.5608<br>(<1e-10)*** | 0.2889<br>(<1e-10)*** | 0.3949<br>(1.4e-8)*** |
| $\sigma^2$ | 0.1537<br>(<1e-10)*** | 0.2017<br>(<1e-10)*** | 0.2634<br>(<1e-10)*** |
| $P(S_t = 1 \mid S_{t-1} = 1)$ | | 0.9880 | |
| $P(S_t = 2 \mid S_{t-1} = 2)$ | | 0.9821 | |
| -ln(Likelihood) | 108.59 | -8.36 | |
| AIC | 0.9919 | -0.0032 | |
| Augmented Dickey-Fuller Test | -1.9168<br>(0.6109) | -4.4297<br>(0.01)*** | |
| Phillips-Perron Test | -10.062<br>(0.5434) | -46.2419<br>(<0.01)*** | |

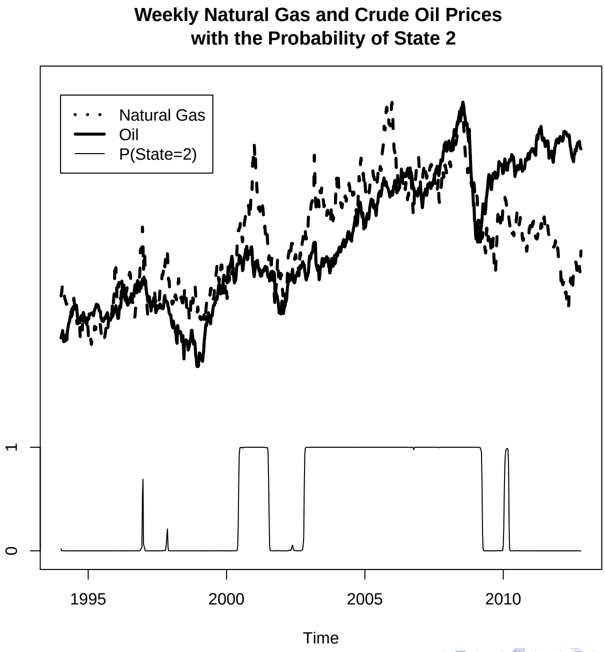
# Results: 2 regimes & weekly prices

**Table 2. Results of Estimating Cointegrating Equation Regressions (equations 1-5) with Weekly Prices**

There are 978 weekly observations for natural gas and oil prices. The one-state model has 975 degrees-of-freedom, and the two-state model has 970. The alternative hypothesis in the augmented Dickey-Fuller and Phillips-Perron tests is to conclude the residual series is stationary. Residuals in the two-state model are weighted by the filtered state probability. p-values are below the coefficient in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

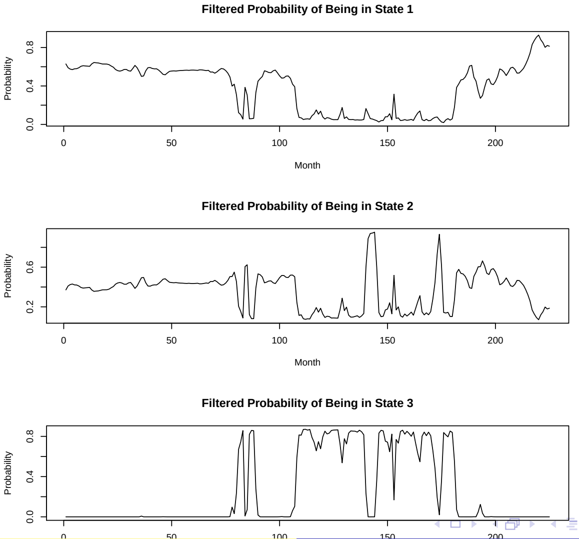
| | **One-State Model** | **Two-State Model** | |
|---|---|---|---|
| State | 1 | 1 | 2 |
| $\beta_0$ | -0.6867<br>(<1e-10)*** | -0.1350<br>(0.0050)*** | 0.2806<br>(0.0043)*** |
| $\beta_1$ | 0.5591<br>(<1e-10)*** | 0.3214<br>(<1e-10)*** | 0.4074<br>(<1e-10)*** |
| $\sigma^2$ | 0.1566<br>(<1e-10)*** | 0.2116<br>(<1e-10)*** | 0.2025<br>(<1e-10)*** |
| $P(S_t = 1 \mid S_{t-1} = 1)$ | | 0.9938 | |
| $P(S_t = 2 \mid S_{t-1} = 2)$ | | 0.9910 | |
| -ln(Likelihood) | 481.2024 | -110.1142 | |
| AIC | 0.9902 | -0.2086 | |
| Augmented Dickey-Fuller Test | -2.3475<br>(0.4312) | -5.2500<br>(0.0100)*** | |
| Phillips-Perron Test | -14.7862<br>(0.2849) | -75.7543<br>(<0.01)*** | |


Monthly Natural Gas and Crude Oil Prices
with the Probability of State 2
Natural Gas
Oil
P(State=2)
1
0
1995
2000
2005
2010
Time

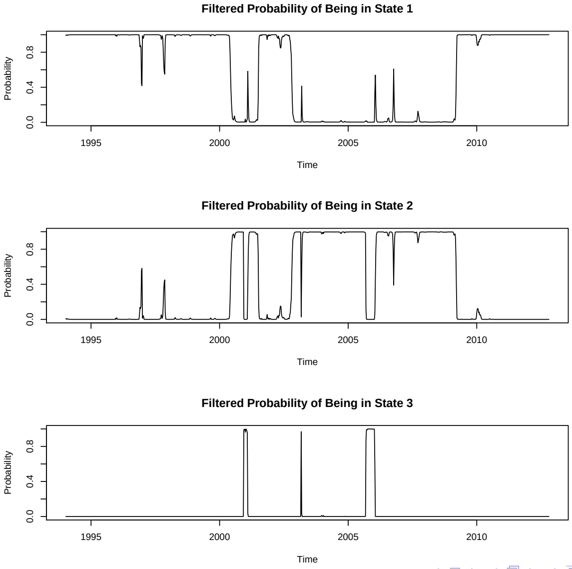


Weekly Natural Gas and Crude Oil Prices
with the Probability of State 2
Natural Gas
Oil
P(State=2)
1
0
1995
2000
2005
2010
Time

# Regime Probabilities: Allowing 3 Regimes, monthly prices



# Regime Probabilities: Allowing 3 Regimes, weekly prices



# Time-Varying Transition Probabilities (TVTP)

We may also wish to allow the transition probabilities to vary according to a set of explanatory variables.

- The probability of transitioning to an electricity price spike regime is likely dependent on reserve capacity and the weather.
- I'll show weather and natural gas storage quantities affect the relative pricing of crude oil and natural gas.

# TVTP: NG and Oil Cointegrating Equation

Let $\mathbf{X}_{t-1}$ be a vector of variables which affect the likelihood of switches between regimes. The cointegrating equation with first-order, two-regime, endogenous Markov switching parameters with time-varying transition probabilities may be written as:

$$P_{HH} = \beta_{0,S_t} + \beta_{1,S_t} P_{WTI} + e_t, \quad e_t \sim N(0, \sigma^2_{S_t})$$

$$p_{ij,t} = P(S_t = j | S_{t-1} = i, \mathbf{X}_{t-1}) = \frac{exp\left(\boldsymbol{\phi}_{ij,0} + \mathbf{X}'_{t-1}\boldsymbol{\phi}_{ij,1}\right)}{1 + exp\left(\boldsymbol{\phi}_{ij,0} + \mathbf{X}'_{t-1}\boldsymbol{\phi}_{ij,1}\right)},$$

$$\forall i, j \in 1, 2 \quad \text{and} \quad \textstyle\sum_{j=1}^{2} p_{ij} = 1$$

$$\beta_{0,S_t} = \beta_{0,1} S_{1t} + \beta_{0,2} S_{2t}$$

$$\beta_{1,S_t} = \beta_{1,1} S_{1t} + \beta_{1,2} S_{2t}$$

$$\sigma_{0,S_t} = \sigma_{0,1} S_{1t} + \sigma_{0,2} S_{2t}$$

where for $m \in 1, 2$, if $S_t = m$, then $S_{mt} = 1$, and $S_{mt} = 0$ otherwise.

# TVTP Example

The variables governing the transition probabilities are:

- the deviation of the number of heating degree days form the long term norm.
- the logged difference of natural gas working gas in storage from its 5-year average.
- An Indicator for Enron's collpase.

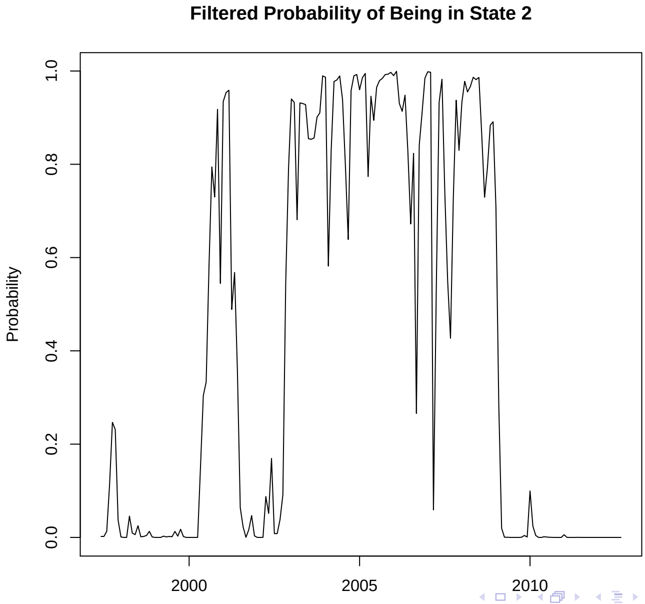
# TVTP: Estimation

## Hamilton Filter: TVTP

```
xi.a[1,] <- (initial.regime.prob*dist[1,])/(o.v%*%(initial.regime.prob*dist[1,]))

for (i in 1:(length(lnoil)-1)){
xi.b[i+1,] <- matrix(c((pnorm(-(g1+g3*VAR1+g4*VAR2))[i]*xi.a[i,1]+
          (1-pnorm(-(g2+g3*VAR1+g4*VAR2)))[i]*xi.a[i,2]), ((1-pnorm(-(g1+g3*VAR1+g4*VAR2)))[i]*xi.a[i,1]+
          (pnorm(-(g2+g3*VAR1+g4*VAR2)))[i]*xi.a[i,2])), nrow=2, ncol=1)
      for (j in 2:length(lnoil)){
      xi.a[j,] <- (xi.b[j,]*dist[j,])/(o.v%*%(xi.b[j,]*dist[j,]))
      model.lik[j] <- o.v%*%(xi.b[j,]*dist[j,])
      }
}
```

We can also maximize the likelihood using the EM algorithm (see Deibold, Lee, and Weinbach (1994)).

# TVTP Example: Filtered Regime Probabilities



# TVTP Example: Parameter Estimates

Table: Parameter estimates of $P_{HH} = \beta_{0,S_t} + \beta_{1,S_t} P_{WTI} + e_{S_t}$ where $S_t \in \{1, 2\}$ and the transition probabilities are time-varying functions of the deviation of the amount of working gas in storage from its 5-year average (STOR) and the deviation of the cumulative number of monthly HDD from its long-term norm (HDD DEV). P-values are below the coefficients in parentheses.

| | Cointegrating Equation | |
|---|---|---|
| | $S_1$ | $S_2$ |
| $\beta_0$ | -0.06572 | 0.07525 |
| | (6.969e-01) | (7.752e-01) |
| $\beta_1$ | 0.34383 | 0.46665 |
| | (8.322e-13)*** | (1.529e-11)*** |
| $\sigma$ | 0.31294 | 0.18495 |
| | (0.000e+00)*** | (0.000e+00)*** |
| | Transition Probabilities | |
| intercept | -0.90138 | -0.45712 |
| | (3.758e-04)*** | (7.656e-02)* |
| HDD DEV | 0.02312 | |
| | (2.452e-10)*** | |
| NG STOR | 1.53150 | |
| | (8.320e-03)*** | |
| Neg Log Likelihood: -40.37859 | | |

# TVTP Example: Parameter Estimates with ENRN Indicator

Table: Parameter estimates of $P_{HH} = \beta_{0,S_t} + \beta_{1,S_t} P_{WTI} + e_{S_t}$ where $S_t \in \{1, 2\}$ and the transition probabilities are time-varying functions of the deviation of the amount of working gas in storage from its 5-year average (STOR), the deviation of the cumulative number of monthly HDD from its long-term norm (HDD DEV), and an indicator for the 6 months after the collapse of Enron (ENRN). P-values are below the coefficients in parentheses.

| | Cointegrating Equation | |
|---|---|---|
| | $S_1$ | $S_2$ |
| $\beta_0$ | -0.2929 | 0.3881 |
| | (3.498e-08)*** | (8.513e-05)*** |
| $\beta_1$ | 0.3716 | 0.3736 |
| | (0.0000)*** | (0.0000)*** |
| $\sigma$ | 0.1813 | 0.1945 |
| | (0.0000)*** | (0.0000)*** |
| | Transition Probabilities | |
| intercept | -1.6204 | -1.5313 |
| | (0.0000)*** | (0.0000)*** |
| HDD DEV | 0.0317 | |
| | (0.0000)*** | |
| NG STOR | 0.7485 | |
| | (0.1027) | |
| ENRN | 1.4220 | |
| | (0.0000)*** | |
| Neg Log Likelihood: -12.6087 | | |

This methodology has recently been found useful for modeling power prices and predicting price spikes. Some relevant research is:

- Mount, Ning, Cai (2006)
- Janczura, Weron (2010)
- among others

We may wish to allow for parameter uncertainty within the model:

> *...a person's uncertainty about the future arises not simply because of future random terms but also because of uncertainty about current parameter values and of the model's ability to link the present to the future. [Harrison and Stevens (1976), pg. 208]*

This can be done by writing the model in state-space form[6] and applying the Kalman filter.

[6] For an introduction to state-space models see Harvey (1989).

# Regime Uncertainty versus Parameter Uncertainty

- In the Markov regime-switching models the regime is the unobservable variable.
- In dealing with parameter uncertainty we first build a structural model of the time series. The model is formulated in terms of unobservable components, which nonetheless have a direct interpretation.
  - We then write the model in state space form (measurement and state transition equations), where the unobservable component is the state[7] vector which evolves according to the state transition equation.
  - We make inference about the state using the Kalman filter.
  - The varying state vector is the set of time-varying parameters.

[7]The term 'state' here is different from 'state' when used in a regime-switching model.

# Testing for Parameter Stability

The Brown, Durbin, and Evans (1975) 'homogeneity test'.

- Null hypothesis: coefficients are equal at all points in time.
- Sample period is divided into $n$ non-overlapping intervals.
- Ratio of 'between groups over within groups' mean sum of squares is the test statistic.
- Test statistic follows an $F$ distribution.

# Code

```
## Brown Durban Evans (1975) 'Homogeneity test' ----
n <- 50
data.test <- cbind(as.vector(data$ng.ret),as.vector(data$stor.2.yr),as.vector(data$hdd.dev),
as.vector(data$cdd.dev))

p <- floor(dim(data.test)[1]/n)

data.test.spl <- array(NA,dim=c(n,dim(data.test)[2],p))
for (i in 1:p){
    data.test.spl[,,i] <- data.test[(1+(i-1)*n):(i*n),]}

## get the residual sum of sqaures for each interval
rss <- rep(NA,p)
for (i in 1:p){
    rss[i] <- t(data.test.spl[,1,i])%*%data.test.spl[,1,i]-t(data.test.spl[,1,i])%*%data.test.spl[,2:4,i]
    %*%solve(t(data.test.spl[,2:4,i])%*%data.test.spl[,2:4,i])%*%t(data.test.spl[,2:4,i])%*%
    data.test.spl[,1,i]
}

## get the full sample rss
rss.full.sam <- t(data.test[,1])%*%data.test[,1]-t(data.test[,1])%*%data.test[,2:4]%*%
solve(t(data.test[,2:4])%*%data.test[,2:4])%*%t(data.test[,2:4])%*%data.test[,1]

## test statistic
k <- dim(data.test)[2]-1
test.stat <- ((dim(data.test)[1]-(k)*p)/((k)*p-(k)))*((rss.full.sam-sum(rss))/sum(rss))

## under Ho distrubuted as F(kp-k,T-kp)
## p-value
1-pf(test.stat, k*p-k, dim(data.test)[1]-k*p)
```

# Kalman Filter Example

- The Kalman Filter is often used in finance to estimate time-varying $\beta$ coefficient in a CAPM type equation[8].
- There are many implementation of the Kalman filter in R and other languages (though you must take care to note what parameters a particular implementation will estimate).
- Through the Kalman filter, the model parameters may be estimated using prediction error decomposition.

[8] The Kalman filter was created by engineers working on control systems. Early applications of the filter were to such things as tracking a missile with a time-series of noisy observations of its location. Because of this, in the first applications of the Kalman filter the parameters were assumed to be known (often by the physics of the situation). Later, creating a likelihood via *prediction error decomposition* was introduced.

# CAPM: Time-Varying $\beta$

Below is an implementation of the CAPM with a time-varying $\beta$ coefficient.

$$r_{s,t} = \alpha + \beta_t r_{m,t} + e_t, \quad e_t \sim i.i.d.N(0, R)$$

$$\beta_{t+1} = \mu + F\beta_t + \nu_t, \quad \nu_t \sim i.i.d.N(0, Q)$$

You can easily allow $\alpha$ to vary with time also. Below however, I have set $\alpha = 0$.

# Kalman Filter Example

## Kalman Filtered Beta Code

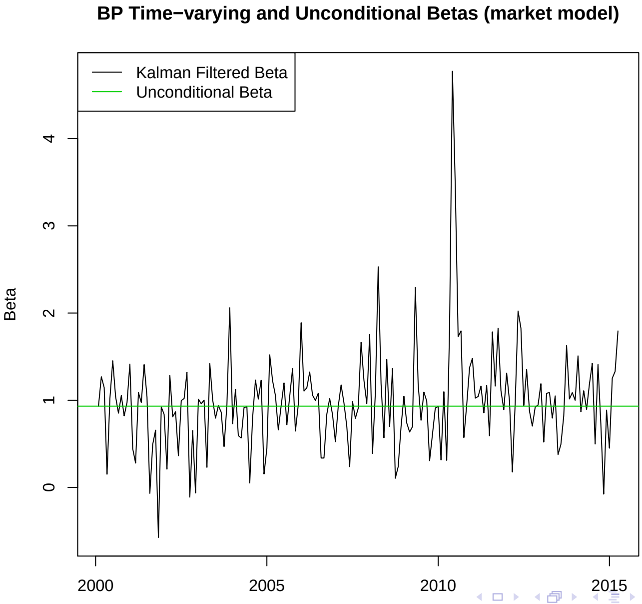
```
lik <- function(theta, market, stock){
    ## R and Q transformed below (squared) -- so they are standard deviations
    R <- theta[1]
    Q <- theta[2]
    F <- theta[3]
    mu <- theta[4]

    beta_tt <- rep(0, length(market))
    beta_tt_1 <- rep(0, length(market))
    eta <- rep(0, length(market))
    f <- rep(0, length(market))
    Ptt <- rep(0, length(market))
    Ptt_1 <- rep(0, length(market))
    ## start at unconditional estimate
    beta_tt[1] <- lm(stock~market)$coef[2]
    Ptt[1] <- (summary(lm(stock~market))$coef[4])^2
    for(i in 2:length(market)){
        ## Prediction ----
        beta_tt_1[i] <- mu + F*beta_tt[i-1]
        Ptt_1[i] <- F*Ptt[i-1]*F+Q^2
        eta[i] <- stock[i]-market[i]*beta_tt_1[i]
        f[i] <- market[i]*Ptt_1[i]*market[i]+R^2
        ## Updating ----
        beta_tt[i] <- beta_tt_1[i]-Ptt_1[i]*market[i]*(1/f[i])*eta[i]
        Ptt[i] <- Ptt_1[i]-Ptt_1[i]*market[i]*(1/f[i])*market[i]*Ptt_1[i]
    }
    logl <- -0.5*sum(log((((2*pi)^length(market))*abs(f))[-1]))-.5*sum(eta*eta*(1/f),na.rm=T)
    return(-logl)
}
```

# Kalman Filter Example

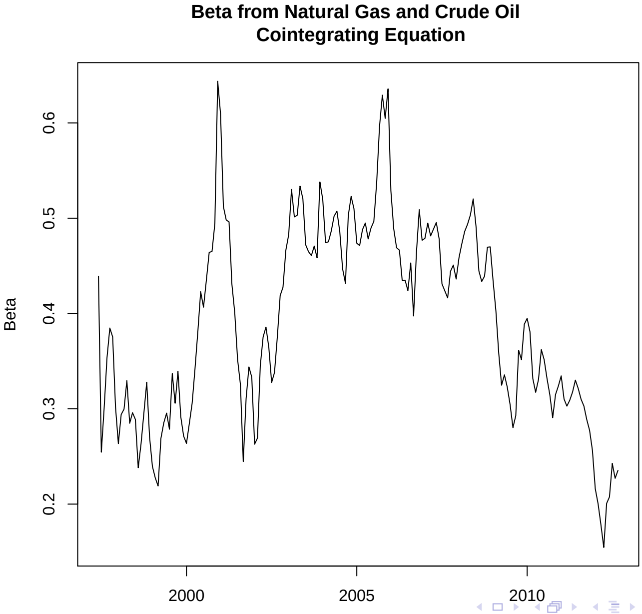
## Kalman Filtered Beta Code Continued

```
theta.start <- c(0.01,0.01, 0.1, 0.1)
max.lik.optim <- optim(theta.start, lik, market=market, stock=stock, hessian=T)

    ## Run though filter to get tv coefficients ----
    R.hat <- max.lik.optim$par[1]
    Q.hat <- max.lik.optim$par[2]
    F.hat <- max.lik.optim$par[3]
    mu.hat <- max.lik.optim$par[4]

    beta_tt <- rep(0, length(market))
    beta_tt_1 <- rep(0, length(market))
    eta <- rep(0, length(market))
    f <- rep(0, length(market))
    Ptt <- rep(0, length(market))
    Ptt_1 <- rep(0, length(market))
    beta_tt[1] <- lm(stock~market)$coef[2]
    Ptt[1] <- (summary(lm(stock~market))$coef[4])^2

    for(i in 2:length(market)){
        ## Prediction ----
        beta_tt_1[i] <- mu.hat + F.hat*beta_tt[i-1]
        Ptt_1[i] <- F.hat*Ptt[i-1]*F.hat+Q.hat^2
        eta[i] <- stock[i]-market[i]*beta_tt_1[i]
        f[i] <- market[i]*Ptt_1[i]*market[i]+R.hat^2
        ## Updating ----
        beta_tt[i] <- beta_tt_1[i]-Ptt_1[i]*market[i]*(1/f[i])*eta[i]
        Ptt[i] <- Ptt_1[i]-Ptt_1[i]*market[i]*(1/f[i])*market[i]*Ptt_1[i]
    }
    logl <- -0.5*sum(log((((2*pi)^length(market))*abs(f))[-1]))-.5*sum(eta*eta*(1/f),na.rm=T)
```

# Time-varying and Unconditional $\beta$: BP



# Time-varying $\beta$ in NG and Oil Cointegrating Equation



# Shiny App and Code

I have put a shiny app where you can input a stock's ticker and time interval, and see a time series of the Kalman filtered beta coefficient, here: `https://mattbrigida.shinyapps.io/kalman_filtered_beta`

Note the app is too slow at this point—I hope to speed it up with Rcpp.

The code is on GitHub here:
`'Matt-Brigida/R-Finance-2015-Kalman_beta'`

# Regime-Switching in State-Space Models

It is straight-forward to include time-varying parameters in a regime-switching model by putting it in state-space form. At this point you can combine the Hamilton filter with the Kalman Filter (Kim (1994)).

- This recognizes the uncertainty in current and future parameter values.
- For example: Does the sensitivity of electricity prices to outages change in an uncertain fashion?

$$\Delta P_t = \mathbf{X}_{t-1}\boldsymbol{\beta}_t + e_t$$
$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t$$
$$\boldsymbol{\eta} \sim N(0, \mathbf{R})$$
$$e_t \sim N(0, \sigma^2_{S_t})$$

# State Space Model with Markov Switching: NG & Oil

To incorporate parameter uncertainty into our two-state Markov regime-switching model of the natural gas and crude oil cointegrating equation, we can write the equation as:

$$P_{HH} = \beta_{0,S_t} + \beta_{1,t,S_t} P_{WTI} + e_t, \quad e_t \sim N(0, \sigma^2_{S_t})$$

$$\beta_{1,t,S_t} = \mu_{S_t} + F_{S_t} \beta_{1,t-1,S_{t-1}} + G_{S_t} v_{S_T}$$

$$P(S_t = j | S_{t-1} = i) = p_{ij}, \ \forall j \in 1, 2, \text{ and} \sum_{j=1}^{2} p_{ij} = 1$$

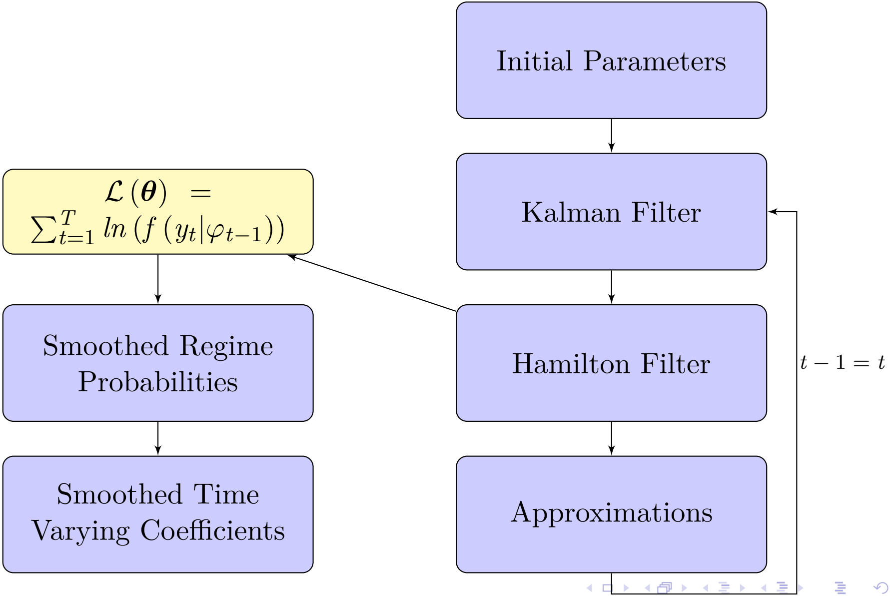
$$\beta_{0,S_t} = \beta_{0,1} S_{1t} + \beta_{0,2} S_{2t}$$

$$\beta_{1,t,S_t} = \beta_{1,t,1} S_{1t} + \beta_{1,t,2} S_{2t}$$

$$\sigma_{0,S_t} = \sigma_{0,1} S_{1t} + \sigma_{0,2} S_{2t}$$

where for $m \in 1, 2$, if $S_t = m$, then $S_{mt} = 1$, and $S_{mt} = 0$ otherwise.

# Overview of Kim (1994) Algorithm

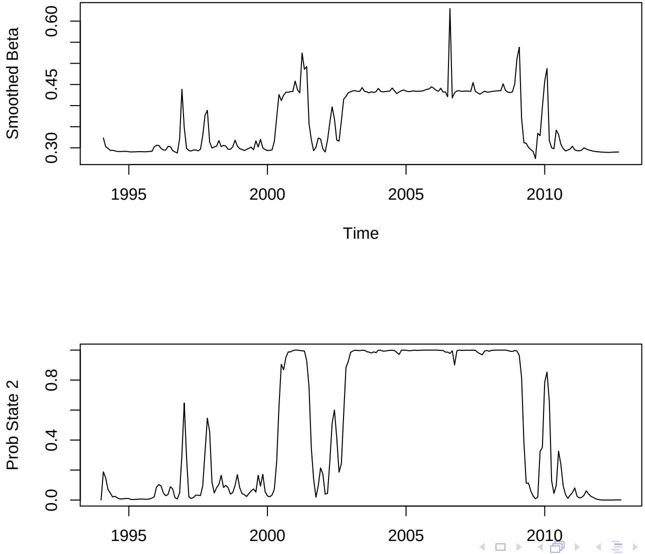


# Overview of Kim (1994) Algorithm

We then take the estimated parameters $\hat{\boldsymbol{\theta}}$ and use them in the smoothing algorithm. The smoothing algorithm will give us:

- Smoothed regime probabilities: $\forall t, j, \quad P[S_t = j|\varphi_T]$
- Smoothed estimates of the time-varying coefficients: $\beta_{t|T}$

# State Space Model with Markov Switching: NG & Oil



# Contact Info

work: mbrigida@clarion.edu
website: http://complete-markets.com
Github: github.com/Matt-Brigida
Twitter: @mattbrigida

# References


- Brown, R. L., J. Durbin, and J. M. Evans (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society* B37, 149–192.
- Diebold, F.X., J.H. Lee, and G.C. Weinbach (1994). Regime switching with time varying transition probabilities. In ‘Nonstationary Time Series Analysis and Cointegration’. Oxford University press.
- Filardo, A.J. (1998). Choosing information variables for transition probabilities in a time-varying transition probability Markov switching model. Federal Reserve Bank of Kansas City research paper.
- Gregory, A.W. and B.E. Hansen (1996). Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics* 70, 99–126.

# References


- Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357-384.
- Harrison, P.J. and C.F. Stevens (1976). Bayesian forecasting. *Journal of the Royal Statistical Society*, Series B, 38, 205–247.
- Harvey, A.C. (1989). Forecasting structural time series models and the Kalman filter. Cambridge: Cambridge University Press.
- Kim, C.J. (1994). Dynamic linear models with Markov-switching. *Journal of Econometrics* 60, 1-22.