

Regime Change

Rob Hayward*

April 12, 2014

Abstract

Hyman Minsky argued that financial system would tend to evolve from a system characterised by *hedge financing* into more fragile regimes of *speculative financing* and *Ponzi financing*. Though this process is not directly observable, there are financial market outcomes that are more likely to be observed in each of these regimes. Analysis of the returns to the attempt to take advantage of deviations from *uncovered interest parity* (UIP), called the *carry-trade*, make it possible to identify these regimes. Return characteristics change as financial instability develops so that the process can be modeled as a Markov chain where the states are unobserved but the outcome that is conditional on the state is used to uncover the parameters. Identifying the financial state allows is used as a means of understanding more about financial instability.

1 The Minsky Model

The Minsky model of endogenously increasing financial fragility has rightly received increased attention in the aftermath of the 2007 - 2008 global financial crisis. The model suggests that a period of economic stability will provide the conditions for a gradual transition of lending within the economy from one characterised as *hedge* though that which is more *speculative* and into something that can be described as Ponzi. Under hedge financing the level of debt will not tend to extend beyond the ability to repay interest and principal; under speculative lending debt-to-equity ratios start to rise and the level of borrowing relative to income is elevated to a point where it becomes difficult to repay principal without asset price appreciation; under the Ponzi

*University of Brighton Business School, Lewes Road, Brighton, BN2 4AT; Telephone 01273 642586. rh49@brighton.ac.uk

lending there is insufficient revenue to repay principal or interest and asset price appreciation or increased borrowing is necessary to maintain the debt. The Ponzi period tends to

While this is a plausible description of rising financial fragility, it is often difficult to determine at what position the economy is in the financial cycle. While measuring debt-to-equity ratios and the scale of bank lending may provide some indication about the regime that is in place, this is imprecise and it is clear that financial services business evolve in ways that make loan counting inadequate. The recent financial crisis showed that innovations like *Collateralised Debt Obligations (CDO)* can allow an increase in debt that does not directly in bank lending.

2 Hidden Markov Models

These models are used extensively in biometrics to analyse genome structure. In particular, they are used to assess the underlying regime that is influencing the gene structure.

[Chapter 10 Biometric Text on HMM](#) has an excellent overview of markov HMM and the R code necessary. One component of this that could be of interest is the assertion in on-line biometrics text that it is a problem to find the underlying state that produced the DNA outcome. The equivalent of this for the crash model is to find the underlying Minsky state that produced the market activity.

This is the folder for building the regime-switching model. The three latent states are the periods of shock, of calm and of crisis. These are unseen but may be identified by other variables (such as the level of international risk aversion (VIX) or the state of domestic political uncertainty (see that buy at Yale???)). Alternatively, it may be possible to draw the states from the data and compare the information that is supplied by the data with that from what is known about political and economic developments at the time. It is also possible to assess the probability that there will be a switch from one regime to another. This is a Markov-switching model.

The motivation for this research is to help to identify regimes and therefore help authorities determine the risk of financial crisis or crash. Use the FX market to get a broader understanding of some of the forces at work.

Some additional motivations are necessary. There is some scope to try copy the analysis of Asian stock market with the identification of different states. This eventually divided the markets into something that looked like developed and less developed. Look into this. It could that fixed and flexible exchange rate regimes would help to explain something about this choice of

exchange rate arrangement. It may also be the case that this can divide the regimes into financially robust and financially fragile.

There are two ways to try to identify the regimes: use external indicators such as the VIX index or other indications of domestic economic or political uncertainty; use the internal structure of the data to identify the periods through which the financial system is evolving. From Maarten's speech (can be deleted later). In a *mixture model*, each observation is assumed to be drawn from a number of distinct sub populations. These can be called *component distributions*. The distribution from which the component is drawn is not immediately observable and is therefore represented as a *latent state*.

A mixture distribution is defined as

$$p(Y_t = y) = \sum_{i=1}^N p(Y_t = y | S_t = i) P(S_t = i) \quad (1)$$

where,

- $S_t \in 1, \dots, N$ denotes the latent state or class of observation t
- $P(S_t = i)$ denotes the probability of the latent state t equals i
- $p(Y_t = y | S_t = i)$ denotes the density of observation of Y_t conditional on latent state being $S_t = i$.

3 Perth Water Example

In the Perth water example that Maarten presented, there are three different models (level, linear and quadratic). The change model assumes that there is one or more discrete change points. It may be the mean, trend or other parameters that change. The transition matrix defines the change points.

$$\begin{pmatrix} p_1 & 1 - p_1 \\ 0 & 1 \end{pmatrix}$$

Where p_1 is the probability that the system will be in state 1. Otherwise, there is a switch to state two. This matrix can be extended for multiple states. What is interesting in the model for Perth water flow is that there is a transition to a second state where there is a lower mean and a lower standard deviation. Therefore, the water level is lower and the variability of the water is less.

Ask Maarten how it determines whether the states can switch back and forth or stay fixed? It looks as if the answer to this question is the transition

matrix can be restricted to ensure that this happens. On page 21 of Maarten's report, it says that, in the EM algorithm, probabilities initialised as 0 or 1 will remain there. However, this does not happen with the bull and bear market example. Why???

There is an example of this method in the analysis of SP 500 regimes: Gains with low volatility and losses with high volatility. The model will identify the most likely

It may be possible to start from the distribution of the returns in the period of crisis and the period of calm and then to move to the identification of transition process. This would be like the slide 9 of Maarten's London R presentation.

In the *dependent mixture model* states are assumed to be statistically dependent. This is consistent with the Minsky theory that the period of calm creates the conditions for the crash. The process underlying the state transitions is a *homogenous first order Markov process* (look this up for additional definition). This process is completely defined by the initial state probabilities.

$$P(S_1 = 1), \dots, P(S_1 = N)$$

and the state transition matrix,

$$\begin{pmatrix} P(S_t = 1|S_{t-1} = 1) & P(S_t = 2|S_{t-1} = 1) & \dots & P(S_t = N|S_{t-1} = 1) \\ P(S_t = 1|S_{t-1} = 2) & P(S_t = 2|S_{t-1} = 2) & \dots & P(S_t = N|S_{t-1} = 2) \\ \vdots & \vdots & \ddots & \vdots \\ P(S_t = 1|S_{t-1} = N) & P(S_t = 2|S_{t-1} = N) & \dots & P(S_t = N|S_{t-1} = N) \end{pmatrix}$$

The models are estimated using the Expectation-Maximisation (EM) or numerical optimisation (when parameters are constrained). The dependent mixture model is made up of three sub models:

1. The prior model: $P(S_1|x, \theta_{prior})$
2. The transition model: $P(S_t|x, S_{t-1}, \theta_{trans})$
3. The response model: $P(Y_t|S_t, x, \theta_{resp})$

See Maarten's notes to see how these are implemented.

3.1 Agn Timmerman (2011)

Looking at how abrupt changes in regime can lead to changes in the way that the system works. The different regimes can be associated with different

underlying distribution of returns. This can allow the understanding of the non-linear and non-normal distribution within normal or linear framework. At the extreme, the regime switch model can incorporate a *jump model* with one change and time-varying parameter models that have a large number of regimes.

The broad framework for the model is to model a discrete state $s_t \in \{0, 1, \dots, k\}$

$$y_t = \mu_{s_t} + \phi_{s_t} y_{t-1} + \sigma_{s_t} \varepsilon_t, \quad \varepsilon_t \sim iid(0, 1) \quad (2)$$

The process governing the underlying regime must also be defined.

$$Pr(s_t = 0 | s_{t-1} = 0) = p_{00} \quad \text{and} \quad Pr(s_t = 1 | s_{t-1} = 1) = p_{11} \quad (3)$$

More generally, the transition could be time-varying and could be dependent on the time spent in the regime. See Durland and McCurdy (1994) for time or Diebold, Lee and Weinbach (1994) for examples where the transition probabilities depend on some other state variables. For example, the interest rate spread. Could the VIX index or other factors be used? Vix index would indicate heightened international tension as one element that affects the probability of transforming from one state to the next. This is very similar to the bubble bursting so any information that improves the ability to identify bubbles bursting would be beneficial.

In this case $p_{ij}(t) = \Phi(z_t)$, where z_t is the conditioning information and Φ could be a logit or probit model.

3.2 Hamilton, Palgrave

From the Hamilton paper for the Palgrave dictionary, the transition matrix will determine the probability of switching from one regime to another. If the transition has $b_{22} = 1$ this implies that the regime switch is permanent. More likely, it will be temporary.

The use of the regime-switch allows the transition from one regime to another to be the result of something that is more than just a deterministic process. There are a number of ways that this model could be expanded. Dueker has a model where the degrees of freedom from a Student-t distribution change with the regime.

4 Literature

The use of the VIX index to identify times of risk aversion can make the period of crisis an observable state of the world. This is about a regime-switching model. This will treat the states of the world to be unobservable

and to fit a regime-switching model (build like that of Hamilton (1989). These notes from Ghysels. There are two states of the world: crisis and moderation. If the system is in a crisis, it stays there with probability p ; it switches to moderation with probability $1 - p$. If in moderation, the system stays there with a probability q and switches to crisis with probability $1 - q$. If the probabilities change over time, there is no longer a *homogenous Markov Chain*. Ghysels has a seasonal dummy for the probabilities that represent the months or quarters.

Can the probabilities change over time? This may be the result of changes in the resilience of the financial system. **Mixture Hidden Markov Models** Hidden models help to calafify the regime under which securities trade. Model takes into account the unobserved hetrogeneity across time. This could be extended to space (for different countries).

Can the system be used to compare to fixed and floating exchange rates.

5 Methods

5.1 Data

The data are a sample of CEE carry trades that have been compiled from raw exchange rate and interest rate data. They show a range of possible carry trades that could have been conducted.

The data are constructed with the function *forp* in the *Raw.R* file. This programme will create a sample of carry-trade profits from the inputs that will provide a funding currency from the set of US dollar, Euro, Swiss Franc and Japanese yen; an investment currency from the set of CEE countries provided; a time period of 1 month or 3 months. Is there any need for any others funding curenecies? I don't think so! It is possible to add other investment currencies. There is ISK, TRY and NOK as reference. It would be possible to have shorter time periods for the carry trade. This would require the addition of the appropriate times series for LIBOR or deposit rates.

The formulra for the carry can be taken from the doctorate.

5.2 Regime selection

There are three states or regimes: the first is caution, the second is build and the third is caution. These three regimes or states loosely correspond to the three stages of financial instability: hedge finance, speculative finance

and Ponzie finance. It is assumed that the Ponzie stage swiftly turns into a crash.

Maybe put the R details in an appendix.

The returns are modeled as a simple model of the level of returns and the standard deviation around this level. It is assumed that there are three different regimes so the level and the standard deviation around this level is allowed to change with the regime.

Three regimes are identified with the 1-month investment in Polish zloty funded by US dollars. These regimes are: caution, carry-build and crash.

5.3 Wikipedia

From wikipedia. **Hidden Markov Models**. The system is a Markov process. There are unobserved, hidden states. The Markov property A Markov chain satisfied the Markov property. That means that it depends only on the current state. It has no memory. *Brownian motion*, for example, is a Markov process. In a Markov chain, the state is observed and the transition probabilities are the only parameter. In a HMM, each state has a probability distribution over the possible outcomes. These can be called *tokens*. A series of *tokens* can tell us something about the states. A hidden Markov model is a generalisation of the mixture model where hidden or latent variables are related through a Markov process. These latent variables control the outcome token. A mixture model is a model with sub-populations where the observations do not allow identification of the sub populations. A mixture model has a mixture distribution. A mixture distribution is a distribution of a random variable that is derived from a number of random variables that are related to each other. If the variable is continuous, this is a mixture density. The *mixture components* are combined together to form the mixture distribution with certain *mixture weights*. Discrete are sometimes called *compound distributions*. There is a difference between adding together two normal distributions (where there will be a new normal with a new mean) and the mixture distribution (where there will be twin peaks).

Mixture Hidden Markov Models The observed response y_{it} is the return of stock market i at time t . There are also two latent variables: a time-constant discrete latent variable and a time-varying discrete latent variable. The former is denoted by $w \in 1, \dots, S$ captures unobserved heterogeneity across markets; $z_t \in 1, 2$ is a two-state, time-varying latent variable.

$f(y_{it}|z_t)$ is assumed to have a multivariate normal density function. This distribution is characterised by $\theta_k = (\mu_k, \sigma_k^2)$. Excluding the states w , there are the initial state probabilities to be determined, the 2 transition probabilities and the conditional mean and conditional variance to be esti-

mated. This is done by Maximum likelihood using the log-likelihood function $l(\varphi, y) = \sum_{i=1}^n \log f(y_i; \varphi)$. This is a problem that can be solved with the *Expectation-Maximization (EM) algorithm*. Dempster, Laird and Rubin (1977). The E step computes the joint conditional distribution of the latent variables given the data and the current provisional estimates of the model parameters. The M step ML methods are used to update the parameters using the estimated densities of the latent variables as weights. For hidden Markov models, a special variant of the EM algorithm is proposed (called the *forward-backward* or *Baum-Welch* algorithm (Baum et al 1970).

The markets are categorised as either *bull* or *bear* markets. This is consistent with the return volatility of markets. About 25% of the returns are in the bear markets. The evidence that the cluster of markets that are closer to emerging markets are more likely to be in a bear market than the developed markets. In addition, it shows that there is persistence. Once a market is in a particular regime (z) it is likely to stay there. Cluster 2 shows a lower propensity to move to a bear market from a bull market. This is consistent with what is known about emerging markets. The picture of the estimated posterior bull market regime is shown for emerging and developed markets. It shows that emerging markets are more likely to suffer bear markets.

The second paper (on rating agencies - included in the "other" folder. Uses the mixture model against the alternative of a pure Markov chain. In the pure Markov chain, the future depends only on the present. However, with the mixture model, the future depends on the past. This means that it is important to know which of the latent sub-groups the firm is in as this will tell you more about the probability of default. This knowledge will be based on the whole sample of ratings. There are A and Q processes. What determines whether the rating evolves according to A or Q? It appears to be partly the result of the industry. The wholesale and retail trades are the most dynamic. There is clearly a heterogeneity that is ignored by the standard Markov model of rating migration.

Notes from the [Leeds notes](#). The aim is to find patterns in time. This uses the example of seaweed and traffic lights. For the traffic lights, there is a state machine where the different states follow each other. Each state is dependent only on the previous state. This is a deterministic system. The weather is not deterministic. There may be three states: wet, cloudy and sunny. The Markov assumption says that the state depends only on the previous state. This is a simplification that makes the problem easier to solve. Some information may be lost with the simplification. The Markov process moves from state to state, depending only on the previous n states. This is called an *order n model*, where n is the number of states affecting the choice of the next state. With the weather example there are 9 possible transformations

(M^2). The probability of each transition is assigned a probability called *state transition probability*. These are collected into a *state transition matrix*. The probabilities do not vary with time. This is a (unrealistic) assumption. To start the system, there is a *vector of initial probabilities*. This is the π vector.

The first order Markov process has three elements:

1. states
2. π vector
3. state transition matrix

Sometimes the Markov model is not sufficient to fully describe the process. In the weather example, the weather may not be observable but seaweed is evident. There is a probabilistic relationship between the seaweed and the weather. More realistic is the identification of hidden states of the mouth through the sounds that can be identified. The observables are related to the hidden states. It is assumed that the hidden states (the weather) are modelled by a simple first order Markov process. The connections between the hidden states and the observable states represent the probability of generating a particular observed state given that the Markov process is in a particular state. This is the *confusion matrix* which gives the probabilities of observable states given a particular hidden state.

The Hidden Markov Model (HMM) is a tripple (π, A, B) where,

1. π Vector of initial state probabilities
2. $A = (a_{ij})$ the state transition matrix $Pr(x_{it}|x_{jt-1})$
3. $B = (b_{ij})$ the confusion matrix $Pr(y_i|x_j)$

Once a system can be described as HMM, three problems can be solved.

1. finding the probability of an observed sequence given a HMM (evaluation)
2. finding the sequence of hidden states that most probably generated the observed sequence (decoding)
3. generating a HMM given a sequence of observed observations (learning)

5.4 Evaluation

There are a number of HMM with sets of (π, A, B) tripples, which HMM geneated the given sequence? For example, there may be 'summer model' and a 'winter model' and it may then be possible to determine the season from the seaweed sequence. The *Forward algorithm* is used to calculate the probability of an observation sequence given a particular HMM and hence the most probable HMM. In speech recognition, the HMM represent different words and the most likely HMM determines the word.

5.5 Decoding

It is most usual to find the hidden states that generated the observed sequence. Finding the hidden states is important because they are not directly observable. A blind hermit may feel the seaweed but cannot see the weather. The *Viterbi algorithm*. This could also be used to determine the syntactic class of words (nouns, verb etc) from the words themselves.

5.6 Learning

This is the most difficult task. Take a set of observations and fit the most probable HMM. The *Forward-backward algorithm* is used when the A and B matrices are not directly (empirically) measurable.

6 Forward Algorithm

With three states, three observations and the parameters of the model known, the aim is to find the most likely hidden sequence. It would be possible to find each possible sequence and sum the probabilities.

7 Results

Three regimes can be identified: number one, the cautious period; number two, the time when carry positions are being built; number three, the crash. This works for PLNUSD, HUF EUR.