A Note on Using Markov Regime Switching Models for Jumps

Jiachuan Tian

Jun.6 2015

While structural break, threshold and asymmetric cointegration models can allow us to characterize the linear and nonlinear dynamics in price transmission in level, it is of equal interest to differentiate across the type of price change to consider what might be thought of a typical price changes as well as extreme price change associated with either structural change or mean reverting as in what we call price jumps.

Jumps in prices are usually modelled as the jump components in the jump-diffusion model in high frequency data. On the other hand, in low frequency data, the mechanism behind a price of "jump" or what we will define as a price spike is not well defined. And there could be several different structures in DGM that could cause jumps or spikes. This paper focuses on modelling jumps and jump transmission in energy and commodity prices. In the paper, I present a model to predict the probability of jumps and jump size conditioning on a set of exogenous variables. In particular, if I choose the exogenous variables to be the indicators of presence of jumps in other price series, then such model can be used to characterize jump transmission.

In time series literature, the concept of "jumps" is related to outlier effects. As Fox(1972), Hillmer et al. (1983), Tsay (1988), Chen and Tiao (1990), Chen and Liu (1993) discussed, there are five types of outlier effects: innovational outliers (IO), additive outliers (AO), level shifts (LS), temporary changes (TC) and variance change (VC). Innovational outliers (IO) refer to those outliers that have effect not only on the particular observation but also subsequent observations, while additive outliers (AO) refer to those that may be mean reverting (zero expectation) and affecting a single observation. Level shifts (LS) and variance change (VC) represent structure changes. Temporary change (TC) corresponds to temporary level shift.

Intuitionally, jumps or spikes refer to those temporary and structural-change-free outliers that have effect on subsequent observations. Thus, jumps should most appropriately be modelled as innovational outliers (IO). To formally define innovational outliers, we follow Chen and Liu and define a price series p_t^i subject to the influence of non-repetitive events as:

$$p_t^i = p_t^{i^*} + \omega^i \frac{A^i(B)}{G^i(B)H^i(B)} \chi\left(t \in \mathcal{J}^{i^*}\right)$$

where *i=1,..,7* for we have seven price series; *t=1,...,T* and *T* is the number of observations for the series; $p_t^{i^*}$ is a latent ARIMA(p, d, q) process, $p_t^{i^*} = (\theta^i(B)/\alpha^i(B)\phi^i(B)){Z_t}^i$and ${Z_t}^i \sim WN(0, {\sigma_z}^2)$; $\theta^i(B)$, $\alpha^i(B)$,$\phi^i(B)$ are polynomials of *B,* the backward shift operator; $\chi(\cdot)$ is the indicator function; $\mathcal{J}^{i^*}$is the set of possible latent time locations of outliers (jumps); and ω^i represents the magnitude of the outlier effects, i.e. jump size; and $A^i(B)/G^i(B)H^i(B)$ denotes the dynamic pattern of the outlier effects. In particular, if $A^i(B)/G^i(B)H^i(B) = \theta^i(B)/\alpha^i(B)\phi^i(B)$, then

$$p_t^i = \frac{\theta^i(B)}{\alpha^i(B)\phi^i(B)}({Z_t}^i + \omega\chi(t \in \mathcal{J}^{i^*}))$$

Here the outliers are innovational outliers. As we discussed, jumps can be seen as innovational outliers. Thus, the dynamics of jumps can be characterized as the model above.

A strand of energy and commodity price literature related to jump (transmission) looks into volatility (transmission). Aizenman and Pinto (2005) point out that volatility in energy prices introduces risk to the economic system and high volatility leads to an overall welfare loss. Serra and Zilberman (2011) such volatility may also spill over directly to soft commodities markets (energy as input in soft commodity), and indirectly through ethanol markets. Thus, the increasing volatility in energy price (and thus in soft commodities price) is a major concern for agricultural producer and agents along the food chain (Balcombe, 2009). Motivated by the need to mitigate the effect of volatility, this paper seeks to study volatility and its transmission between energy and soft commodities prices.

While volatility plays an important role in the economy, the concept of volatility is not well defined in literature. Volatility by its meaning should include two kinds of effects: systematic variance changes (VC) and unsystematic oscillations that cannot be characterized by VC, like jumps and spikes. Thus to fully understand volatility, one has to consider all kinds of oscillation effects, especially jumps.

On the other hand, in many works of literature, the word "volatility" is referred to systematic variance changes. In particular, price volatility is usually characterized as the volatility term in continuous time models (for example, high-frequency stock prices as geometric Brownian motion in Black-Scholes model), or heteroskedasticity error structure in discrete time models (for example, low-frequency energy prices as in time series models). In particular, to characterize "volatility" (variance change effect) transmission in energy and commodity prices (low-frequency data), most existing literature use multivariate general auto-regressive conditional heteroskedasticity (MGARCH) model. Harri and Hudson (2009) employ a MGARCH and Chung and Ng test to investigate "volatility" transmission between crude oil futures and a set of soft commodities. Serra and Zilberman (2009) employ Baba-Engle-Kraft-Kroner (BEKK) MGARCH to model price "volatility" in ethanol market. Zhang et al.(2008) also employs BEKK model to investigate "volatility" of US gasoline prices.

Essentially the MGARCH approach is modelling variance change (VC) effects. And as discussed above, variance changes are neither jumps nor the only source of volatility. Thus the MGARCH models cannot be used to model jumps, and since jumps contributes heavily to volatility, these models cannot efficiently characterize volatility transmission with the presence of jumps.

Another approach is to use a **Markov regime switching model**. Higgs and Worthington (2008) use this method to model the Australian wholesale spot electricity markets. Following this approach, I decompose price at time *t* as deterministic and stochastic components. Using previous notation, we have

$$p_t^i = D_t^i + X_t^i$$

where D_t^i is the deterministic component and X_t^i is the stochastic component. Here D_t^i can be modeled as an ARIMA *(p, d, q)* process, and X_t^i is characterized by Markov regime switching model. In particular, define the jump process J_t^i as

$$J_t^i \equiv \chi\left(t \in \mathcal{J}^{i^*}\right)$$

In addition, to illustrate the instantaneous mean reverting behavior, I define $J_t^i = -1$ if there is mean reversion at *t*. Note using this notation, $\mathcal{J}^{i^*} = \{t = 1, \dots, T | J_t^i = 1\}$. And let X_t^i be such that

$$\nabla X_t^i = \begin{cases} \mu_1^i + \sigma_1^i Z_t & \text{if } J_t^i = 1 \\ -\mu_{-1}^i X_{t-1}^i + \sigma_{-1}^i Z_t & \text{if } J_t^i = -1 \\ -\mu_0^i X_{t-1}^i + \sigma_0^i Z_t & \text{if } J_t^i = 0 \end{cases}$$

where $\nabla = 1 - B$ is the first differenced operator. Following Higgs and Worthington (2008), I assume:
A1. A jump is followed by an immediate mean reversion in next period (the duration of jump is 0);
A2. Price will behave "normally" in the next period of a mean reversion (there is no continuous oscillation). If we define the transition probability $\pi(j, k)$ as

$$\pi^i(j, k) \equiv P\left(J_{t+1}^i = j \middle| J_t^i = k\right), \forall t$$

then the two assumptions (A1, A2) can be written as $\pi^i(-1,1) = 1$; $\pi^i(0, -1) = 1$. Combining with other relations, all transition probabilities $\pi(j, k)$ for $\forall k$, $\forall j$ can be calculated using $\pi^i(0,0)$. Markov regime switching models usually assume this probability to be fixed, for example $\pi^i(0,0)$ can be chosen as $\exp(v^i)/(1 + \exp(v^i))$ for some $v^i \in \mathbb{R}$ a parameter.

It is worth noting that the Markov regime switching approach assumes that the transition probability is fixed, thus the probability of the presence of a jump is homogenous for all *t*. This is not a realistic assumption because the dynamics of price would most probably change over the sample period and hence the probability of a jump will change. It is also not easy to use exogenous (economic) variables to predict the probability of a jump in this scheme since the probability of a jump is assumed to be fixed. If I am interested in predicting the probability of a jump given a set of exogenous economic variables, I can model $\pi^i(0,0)$ as a function of these variables. Since most economic variables are not stationary, $\pi^i(0,0)$ as a function of these variables is not stationary. That directly violates the definition of $\pi^i(0,0)$ as a transition probability in a Markov Chain. Thus the Markov regime switching approach, because of its Markov Chain structure, is inherently unable to predict probability of a jump conditioning on a set of exogenous variables.

A third approach is to use autoregressive conditional hazard(ACH) function. Christensen et al. (2012) employ this model to forecast spikes in electricity prices. Following their methods, for some positive integer $T^* \le T-1$, I can model jumps $\{J_t^i\}_{t=1,..,T^*,T^*+1}$ using $\{P_t^i\}_{t=1,...,T^*}$ as a generalized Poisson process conditioning on its own history and other exogenous variables. In particular, for discrete time prices in this research, with J_t^i define as above, I define the counting process $N(T^*)^i$ of number of jumps in time spots $\{1,..,T^*\}$ as $N(T^*)^i \equiv \sum_{t=1}^{T^*} \chi\left(J_t^i > 0\right) \le \left|\mathcal{J}^{i^*}\right|$, where $|\cdot|$ reports cardinality. Let S_j^i be the time of the *j-th* jump in process$\{P_t^i\}_{t=1,...,T^*}$, then I define the filtration $\mathcal{H}_{T^*}^i$ with respect to S_j^i as

$$\mathcal{H}_{T^*}^i \equiv \sigma(S_1^i, \dots, S_{N(T^*)^i}^i) \subseteq \sigma\left(p_1^i, \dots, p_{T^*}^i\right)$$

Here $\mathcal{H}_{T^*}^i$represents the history of jumps in price p_t^i up to time *t*. Then following Hamilton and Jorda (2002), I define the ACH function can be defined as

$$h_{T^*+1}^i \equiv P(J_{T^*+1}^i > J_{T^*}^i | \mathcal{H}_{T^*}^i)$$

In further, define $\psi_{N(T^*)^i+1}^i \equiv \mathbb{E}(S_{N(T^*)^i+1}^i - S_{N(T^*)^i}^i | \mathcal{H}_{T^*}^i)$ as conditional expectation of next interval time. It can be modelled as an ARMA (p, q) process (Fernandes and Gramming, 2006):

$$(\psi_{N(T^*)^i+1}^i)^v = \sum_{j=1}^{p} \alpha_j^i \,(S_{N(T^*)^i+1-j}^i - S_{N(T^*)^i-j}^i)^v + \sum_{j=1}^{q} \beta_j^i \,(\psi_{N(T^*)^i+1-j}^i)^v$$

Hamilton and Jorda (2002) showed that in discrete time model, the ACH function can be expressed using the ARMA conditional duration. That is,

$$h_{T^*+1}^i = (\psi_{N(T^*)^i+1}^i)^{-1}$$

This formula can be in further augmented to include a set of exogenous variables denoted as a vector $Y_{N(T^*)^i+1}^i$, then the ACH function becomes to

$$h_{T^*+1}^i = (\Lambda(\exp\left(-\gamma^{i\prime} Y_{N(T^*)^i+1}^i\right) + \psi_{N(T^*)^i+1}^i)^{-1}$$

where Λ is a link function such that $h_{T^*+1}^i$ is a probability (see Hamilton and Jorda (2002) for specific function form. Let $\theta^i \equiv (\gamma^i, \alpha_1^i, \dots, \alpha_p^i, \beta_1^i, \dots, \beta_q^i, v)$ denote all the parameters for each price series p_t^i, then the log-likelihood function for a sample of *T* intervals is

$$\log L\left(\theta^i\right) = \sum_{t=1}^{T} \{\chi\left(J_t^i = 1\right) log\, h_t^i + \chi\left(J_t^i = 0\right) \log(1 - h_t^i)\}$$

which may be maximized to obtain MLE of the parameters θ^i for each price series p_t^i.

This approach gives a reasonable modelling on the probability of jumps given a set of exogenous variables. On the other hand, it cannot model jump size conditioning on that there is a jump. This paper proposes a new approach to simultaneously model the probability of a jump and jump size conditioning on a set of exogenous (economic) variables using Heckman selection model. In particular, for a given price series p_t^i, to test for jumps (innovational outlier), I assume the DGM of p_t^i is *ARIMA (p, d, q)* with jumps:

$$p_t^i = \frac{\theta^i(B)}{\alpha^i(B)\phi^i(B)}({Z_t}^i + \omega^i\chi(t \in \mathcal{J}^{i^*}))$$

where ${Z_t}^i \sim WN(0, {\sigma_z}^2)$. Use the test proposed by Chen and Liu (1993) (CL test thereafter), I identify and compose an estimate of the set of jumps $\mathcal{J}^{i^*}$. That is,

$$\widetilde{\mathcal{J}^{i*}} \equiv \{t = 1, \dots, T | \text{CL test concludes there is a jump at t}\}$$

It is worth noting that there could be many factors that cause jumps in prices. In stock market, the usual drivers are liquidity (common measures are bid-ask spread, time between two sequential transaction, etc.), macroeconomics news and firm news. It is worth noting that here liquidity measures the easiness of making transaction in stock market. Such concept of liquidity can be generalized to accommodate our case of low-frequency energy and commodity price data. In our case, the measure of liquidity is inventory. When the inventory level of some particular commodity becomes critically low, producers are facing a high risk of stock out. Facing such risk of stock-out, producer would pay much more (than usual) up to the cost when the production is forced to stop. As a result, the low inventory level results in a jump increase in prices. Thus the process of jumps $\{J_t^i\}_{t=1,\dots,T}$ is driven by a set of exogenous variables $\{Y_t^i\}$. Define the observed jump process as $\widetilde{J_t^i} \equiv \chi(t \in \widetilde{\mathcal{J}^{i*}})$. Following Heckman (1979), the latent process underlying the jumps process $\{\widetilde{J_t^i}\}_{t=1,\dots,T}$ is

$$\mathbb{E}(\overline{J_t^i}|Y_t^i) = {Y_t^i}^T\delta + v_{2t}$$

where $v_{2t} \sim N(0,1)$. Define $\widetilde{J_t^i} = \chi(\overline{J_t^i} \geq 0)$, then the probability of a jump can be model as

$$P(\widetilde{J_t^i} = 1 | Y_t^i) = P(\overline{J_t^i} \geq 0 | Y_t^i) = \Phi({Y_t^i}^T\delta)$$

If I choose $Y_t^i = \{\widetilde{J_s^j}\}_{s \leq t}$, that is, the presence of jump in other series to be such exogenous variable, then I have

$$P(\widetilde{J_t^i} = 1 | Y_t^i) = P\left(\widetilde{J_t^i} = 1 \middle| \{\widetilde{J_s^j}\}_{s \leq t}\right) = \Phi(\sum_{s \leq t} \widetilde{J_s^j}\delta_s + v_{2t}), \forall j \neq i$$

This is a jump transmission model. In addition, to model the jump size ω^i, consider a latent variable ω^{i^*} for ω^i depending on a set of exogenous variables:

$$\mathbb{E}(\omega^{i^*}|W_t^i) = {W_t^i}^T\eta + v_{1t} \text{ and } \omega^i = \omega^{i^*}\chi(\widetilde{J_t^i} = 1)$$

Then I have

$$\mathbb{E}\,(\omega^i|W_t^i, \widetilde{J_t^i} = 1) = {W_t^i}^T\eta + \mathbb{E}(v_{1t}|\widetilde{J_t^i} = 1)$$

Note there is interdependence between ω^i and W_t^i, and hence interdependence between ω^{i^*} and $\overline{J_t^i}$. Such interdependence is represented by the correlation in the joint normal distribution of (v_1, v_2) in the Heckman model. In particular, I assume

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1 \\ \rho\sigma_1 & 1 \end{pmatrix}\right].$$

Then the equation above can be written as

$$\mathbb{E}\left(\omega^i\middle|W_t^i, \widetilde{J_t^i} = 1\right) = {W_t^i}^T\eta + \rho\sigma_1\frac{\phi({Y_t^i}^T\delta)}{\Phi({Y_t^i}^T\delta)}$$

This finishes the modelling of jump size ω^i.

As shown above, for chapter 3, I have finished the modelling part. Compared to Christensen et al.(2012), this approach models jumps as a component of the price series rather than a standalone generalized Poisson process. This gives us the flexibility to model the jump size conditioning on the presence of jumps through the consideration of the selection problem inherent in the study of jumps. That is, intuitively, we can only estimate the jump size if there is a jump. The next steps are to first identify the set of exogenous variables, and then estimate the model using data. A possible concern is that this model use a two-step method, thus the estimation methods of the selection model may need some modifications. A most appropriate candidate is to use Bayesian methods to estimate this hierarchical model.