**Purdue University**

**Computational Finance Program**

**MGMT 616 “Asset pricing”**

**term paper**

# GMM parameter estimation for the Double Exponential Jump-Duffusion Process

**by**

**Nik Tuzov**

**ntuzov@purdue.edu**

**December 2006**

## Abstract

The goal of this term paper is to investigate empirically the performance of GMM in estimating the parameters of the double exponential jump-diffusion model (DEJD) proposed by Kou (2002) and Ramezani & Zeng (1998). Ait-Sahalia (2003) showed that, for Merton jump-diffusion model, MLE is preferable to GMM in terms of efficiency. Although no similar study on DEJD has been found, it is reasonable to guess that similar efficiency results hold for DEJD also.
In practical terms, however, GMM is much simpler to implement than MLE; besides, GMM doesn't have the issue of the maximum likelihood function being possibly unbounded.
In this study, we try to estimate DEJD parameters for a particular dataset via GMM / ITGMM in SAS. To obtain the moment conditions, MGF and cumulant–based approaches are used.
In either case, the results are unsatisfactory because of high correlation of sample moments used in GMM.



**Keywords:** Jump-diffusion model, Kou, Ramezani, Zeng, Merton, Double Exponential, Pareto-Beta, GMM, MLE, moment conditions, cumulats, MGF, convergence, derivative pricing.

## 1. Introduction: PBJD and DEJD models

Jump diffusion processes are processes of the form

$$X_t = \sigma W_t + \mu t + \sum_{i=1}^{N_t} Y_i; \quad X_0 = 0.$$

Here $\{W_t; t \geq 0\}$ is a standard Brownian motion, $\{N_t; t \geq 0\}$ is a Poisson process with rate $\lambda$, constants $\mu$ and $\sigma$ are the drift and volatility of the diffusion part, and the jump sizes $Y_i$ are i.i.d. random variables.

$X_t$ is used to model the log-return process as $X_t = \log(S_t / S_0)$ where $S_t$ is the asset price at time *t*.

In the Double Exponential Jump Diffusion Model [3] $Y_i$ has the double exponential density

$$f_Y(y) = p \cdot \eta_1 \cdot e^{-\eta_1 y} \cdot 1_{\{y \geq 0\}} + (1-p) \cdot \eta_2 \cdot e^{\eta_2 y} 1_{\{y \leq 0\}}; \ \eta_1 > 1, \ \eta_2 > 0$$

Parameters $\mu$ and $\sigma$ can be interpreted as the mean and std of one-period log-returns in the absence of jumps. Parameter $p$ is the probability of an upward jump and $\lambda$ is the average number of jumps per one period in both directions. Parameters $\eta_1$ and $\eta_2$ can be interpreted as follows: if the price change occurred instantly just because of a jump, then

$$E[S_t / S_{t-}] = p \cdot \frac{\eta_1}{\eta_1 - 1} + (1-p) \cdot \frac{\eta_2}{\eta_2 + 1}$$

and we can compute an average jump size in % of the "pre-jump" price $S_{t-}$ as

$$( \ E[S_t / S_{t-}] \ - \ 1 \ ) \cdot 100\%$$

Thus, an average upward jump will have a magnitude of $(\frac{1}{\eta_1 - 1}) \cdot 100\%$ and an average downward jump will have a magnitude of $(\frac{-1}{\eta_2 + 1}) \cdot 100\%$.

In the Pareto-Beta Jump-Diffusion (PBJD) model [2] the jump component is generated by two independent Poisson processes. The first process is responsible for upward jumps whose magnitude has the Pareto distribution, and the second process generates downward jumps whose magnitude has the Beta distribution. Both models have six parameters, and the parameters of one model can be easily recovered from the other [6]. Hence, there's no difference between PBJD and DEJD from the estimation perspective. Once the model parameters are estimated, we can use them for derivative pricing as described in [3], [5] and [7] where the parameters are assumed known.

## 2. Estimation: GMM versus MLE

Ait-Sahalia [3] provides some insight on which of the two estimation methods is preferable for Merton Jump-Diffusion model. It turns out that MLE is more efficient than GMM which gives us grounds to believe that similar efficiency results hold for DEJD model also.

However, MLE has a couple of practical disadvantages compared to GMM. First, it is harder to implement than GMM: in [6], MLE involves estimating double improper integrals and it takes 4-7 hours to obtain the first set of parameter estimates for 1256 observations. Second, one has to be careful as to the admissible region for the parameters lest the likelihood function become infinite [1].

In this study, the plan is to estimate the model parameters for a particular underlying asset via GMM and then use the results to estimate the prices for a number of derivatives for that underlying. Then we can compare the observed and estimated derivative prices which can serve as empirical evidence of how good the performance of GMM is. For instance, the mispricing error being no greater than the derivative bid-ask spread can serve as empirical evidence that at least for this particular dataset GMM performance is adequate.

## 3. GMM estimation: cumulant and MGF – based.

Ramezani and Zeng [2] provide and explicit estimation method based on matching the first six cumulants for log-returns. If we use the cumulants as moment conditions in GMM we can get both the point estimates and standard errors for the parameters.

Alternatively, one can try using the moment-generating function of $X_t$ from [5]

$$\phi(\theta,t) = E[e^{\theta X_t}] = \exp(G(\theta)t),$$

$$G(\theta) = \theta\mu + \frac{1}{2}\theta^2\sigma^2 + \lambda\left( \frac{p\eta_1}{\eta_1-\theta} + \frac{(1-p)\eta_2}{\eta_2+\theta} - 1 \right)$$

One can set $t = 1$ and match the sample and population moments of $e^{\theta X_1}$ for a range of $\theta$ values via GMM.

## 4. GMM for FTSE100 data

To implement the plan above we start with FTSE100 index daily data for the period 04/03/06 – 12/20/06.

To get the initial values for GMM, it’s natural to apply cumulant-matching as in [2] (with parameters in PBJD form). It starts with solving a quadratic equation, but in this case the roots turn out complex which is probably caused by inaccurate estimation of higher-order cumulants. Still, reasonable initial values can be obtained through the physical interpretation of the parameters from Part1.

However, iterative GMM procedure in SAS produced negative estimates for jump intensities, and, most importantly, it appears that three out of six moment conditions are identified by SAS as redundant which means that the six moment are highly correlated:

Nonlinear ITGMM Parameter Estimates

| Parameter | Estimate | Approx Std Err | t Value | Approx Pr > \|t\| |
|---|---|---|---|---|
| mu | 0.00021 | 0.000631 | <------ | Biased |
| sigma | 1.343E-6 | 0.000995 | <------ | Biased |
| **lambda_u** | **-1.32E-8** | **0** | **<------** | **Biased** |
| **lambda_d** | **-1.91E-9** | **0** | **<------** | **Biased** |
| eta1 | 129.2029 | 0 | <------ | Biased |
| eta2 | 130.7454 | 0 | <------ | Biased |

WARNING: The covariance across equations (the S matrix) is singular. A generalized inverse was computed by setting to zero the part of the S matrix for the following 3 equations whose residuals are linearly dependent with residuals from earlier equations: h4 h5 h6

Therefore, we try an alternative way and apply GMM using MGF matching with six values of $\theta \in \{-4, \ -2.5, \ -1, \ 1, \ 2.5, \ 4\}$. Because

$$e^{\theta X_1} = \left( \frac{S_{t+1}}{S_t} \right)^{\theta}$$

where $\left( \frac{S_{t+1}}{S_t} \right)$ is a one-period (non-log) asset return, we can match the sample moments of the right-hand-side to the population moments given by MGF (with parameters in DEJD form). The range for $\theta$ is chosen in this particular way since it is generally acknowledged that the sample moments of order higher than four are poor estimates of the corresponding population moments.

This approach, however, doesn't work either. The estimation difficulties are similar to the case of cumulant-matching: the moments are highly correlated which prevents convergence and the upward jump probability estimate turns out negative:

Nonlinear 2SLS Parameter Estimates (Not Converged)

| Parameter | Estimate | Approx Std Err | t Value | Approx Pr > \|t\| |
|---|---|---|---|---|
| mu | 0.00047 | 0.6894 | 0.00 | 0.9995 |
| sigma | 0.00825 | 37.1224 | 0.00 | 0.9998 |
| lambda | 0.012214 | 2059.1 | 0.00 | 1.0000 |
| **p** | **-0.35069** | **141962** | **-0.00** | **1.0000** |
| eta1 | 80.75856 | 196296 | 0.00 | 0.9997 |
| eta2 | 79.94541 | 275995 | 0.00 | 0.9998 |

The sample correlation matrix computed for the six moment conditions under the initial parameter values is as follows:

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 0.9999 | 0.9997 | -0.9993 | -0.9988 | -0.9981 |
| 0.9999 | 1.0000 | 0.9999 | -0.9996 | -0.9993 | -0.9988 |
| 0.9997 | 0.9999 | 1.0000 | -0.9999 | -0.9996 | -0.9993 |
| -0.9993 | -0.9996 | -0.9999 | 1.0000 | 0.9999 | 0.9997 |
| -0.9988 | -0.9993 | -0.9996 | 0.9999 | 1.0000 | 0.9999 |
| -0.9981 | -0.9988 | -0.9993 | 0.9997 | 0.9999 | 1.0000 |

It now clear why SAS has difficulties estimating the model. The underlying cause for such a high correlation is that $e^{\theta X_1} \approx 1 + \theta X_1$ for small values of $\theta X_1$ which is indeed the case because the average daily log-return $X_1$ is close to zero. Hence, all six moment conditions in GMM become almost identical to each other.

Ait-Sahalia [4] managed to avoid this problem by using absolute centered moments for Merton jump-diffusion model. However, the population absolute central moments for DEJD model are unknown and that approach is not available.

Our next step is to try a larger time scale hoping that since an average weekly log-return is larger, the correlation problem may be mitigated. At the same time, one can say that according to the empirical results in [6], if DEJD has an edge over simpler models (Merton or plain GBM) at all, it is likely to be present on a scale that is no coarser than daily. The coarser the scale, the larger the chance that it is not necessary to include an elaborate jump component or any jump component, for that matter.

Although for weekly data (08/05 – 12/06) cumulant-matching does not produce complex roots, the SAS GMM still fails to converge:

Nonlinear ITGMM Parameter Estimates (Not Converged)

| Parameter | Estimate | Approx Std Err | t Value | Approx Pr > \|t\| |
|---|---|---|---|---|
| mu | 0.002452 | 0.00168 | <------ | Biased |
| sigma | 0.000219 | 0.000019 | <------ | Biased |
| lambda_u | -5.81E-9 | 1.011E-9 | <------ | Biased |
| lambda_d | 1.734E-6 | 0.000011 | <------ | Biased |
| eta1 | 392.9063 | 0.0953 | <------ | Biased |
| **eta2** | **-77220.4** | 0 | <------ | Biased |

NOTE: The model was singular. Some estimates are marked 'Biased'.

Unfortunately, MGF – based GMM produces a similar picture:

Nonlinear 2SLS Parameter Estimates (Not Converged)

| Parameter | Estimate | Approx Std Err | t Value | Approx Pr > \|t\| |
|---|---|---|---|---|
| mu | 0.004474 | 0.6577 | <------ | Biased |
| sigma | 0.01406 | 4.3543 | <------ | Biased |
| lambda | 0.118455 | 0 | <------ | Biased |
| p | -0.00527 | 274.1 | <------ | Biased |
| eta1 | 32.01605 | 180436 | <------ | Biased |
| eta2 | 50.96539 | 23049.6 | <------ | Biased |

NOTE: The model was singular. Some estimates are marked 'Biased'.

The sample correlation matrix computed for the six moment conditions under the initial parameter values is as follows:

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 0.9996 | 0.9983 | -0.9952 | -0.9916 | -0.9870 |
| 0.9996 | 1.0000 | 0.9996 | -0.9976 | -0.9949 | -0.9912 |
| 0.9983 | 0.9996 | 1.0000 | -0.9992 | -0.9974 | -0.9947 |
| -0.9952 | -0.9976 | -0.9992 | 1.0000 | 0.9995 | 0.9980 |
| -0.9916 | -0.9949 | -0.9974 | 0.9995 | 1.0000 | 0.9995 |
| -0.9870 | -0.9912 | -0.9947 | 0.9980 | 0.9995 | 1.0000 |

It looks slightly better than that on a daily scale, but still SAS identifies at least one moment out of six as redundant in GMM and there's no convergence.

The original plan was to use the estimated parameters to price a number of European options on FTSE100 and then to see how large the mispricing error is, but because of failed convergence it doesn't make sense.

## 5. Final remarks.

In this small study an attempt was made to avoid a computationally challenging MLE method and obtain reasonable parameter estimates for DEJD/PBJD model via GMM procedure that appeared to be a lot more straightforward.

Although the attempt turned out unsuccessful, its failure is not inherent to GMM methodology because if one manages to specify the six moment conditions that are more orthogonal to each other than those used in this study, GMM is expected to produce adequate results. In the meantime, practitioners have to recourse to MLE [6].

## References

[1] Honore (1998) “Pitfalls in estimating Jump-Diffusion Models”

[2] Ramezani, Zeng (1998) “Maximum Likelihood Estimation of asymmetric Jump-Diffusion Processes: application to security prices”

[3] Kou (2002) “A Jump-Diffusion model for option pricing”

[4] Ait-Sahalia (2003) “Disentangling diffusion from jumps”

[5] Kou, Wang (2003) “First passage times of a jump diffusion process”

[6] Ramezani, Zeng (2004) “An empirical assessment of the double exponential jump-diffusion process”

[7] Kou, Wang (2004) “Option pricing under a double exponential jump diffusion model”