1 Equivalent formulations

1.1 Double chain representation

We proceed by defining a new first order homogeneous Markov chain $S_{\omega,Z}^t$ with state space \{1, \ldots, JK\}. The Markov chain $S_{\omega}^t$ is defined by combining the Markov chain $S_{\omega}^t$ and the integer random variable $Z_t$. In addition, let $\Omega = [\omega_{j,k}]$ be a $J \times K$ matrix containing the mixture probabilities and let $U$ and $\iota$ be a $K \times K$ matrix of ones and a $JK \times 1$ vector of ones, respectively. The transition probability matrix of $\Gamma_{\omega,Z}$ is given by

$$\Gamma_{\omega,Z} = \iota \text{vec}(\Omega') \odot (\Gamma^\omega \otimes U),$$

(1)

where $\otimes$ and $\odot$ denote the Kronecker and hadamard products, respectively. The initial distribution of $S_{\omega,Z}^t$ has generic element $\omega_{j,k} \delta^\omega$. By incorporating $Z_t$ in $S_{\omega}^t$ via the new Markov chain $S_{\omega,Z}^t$, we obtain the equivalent model representation reported in Figure 1. The term $S_{\omega,Z}^t$ is still a homogeneous first order Markov chain; however, its state space has been enlarged compared to that of $S_{\omega}^t$, and its transition probability matrix has a constrained structure provided by (1). The constrained structure imposed by the transition probability matrix $S_{\omega,Z}^t$ corresponds to a specific ordering of the conditional densities of $Y_t | S_{\omega,Z}^t$.

Figure 1: The model path diagram where $Z_t$ is incorporated in $S_{\omega}^t$, resulting in $S_{\omega,Z}^t$. 

---

*Department of Economics and Business Economics, Aarhus University and CREATES, Fuglesangs Allé 4, 8210, Aarhus V, Denmark. leopoldo.catania@econ.au.dk.
†Department of Economics and Business, University of Catania, Italy. roberto.dimari@unict.it.
‡Department of Economics and Finance, LUISS "Guido Carli" University. Viale Romania 32, 00197 Roma, Italy. CREATE, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. smagistris@luiss.it.
Specifically, let \( j \in \{1, \ldots, J\} \) and \( k \in \{1, \ldots, K\} \) be two indexes spanning the state space of \( S^\omega_t \) and \( Z_t \), respectively. The state space of \( S_{t}^{\omega,Z} \) can be represented as \( \{11, 12, \ldots, 1K, 21, \ldots, jk, \ldots, JK\} \). The conditional densities of the first \( K \) regimes of the \( S_{t}^{\omega,Z} \) Markov chain are the \( K \) mixture components of the first regime of the Markov chain \( S_t^\omega \), from \( K + 1 \) to \( 2K \) of the second regime, until \( (J - 1)K + 1 \) to \( JK \).

### 1.2 Single chain representation

By combining the two independent Markov chains \( S^\omega_t \) and \( S_t^\kappa \) in a third Markov chain \( S_{t}^{\omega,Z,\kappa} \) with state space \( \{1, \ldots, JKL\} \) and transition probability matrix \( \Gamma^{\omega,Z,\kappa} = \Gamma^\kappa \otimes \Gamma^\omega \), we obtain a representation of the DyMiSk in terms of a single Markov chain, with generic element of the initial distribution given by \( \omega_{t,j,k} \delta_j \delta_{\kappa} \). As for the previous representation, the state space of \( S_{t}^{\omega,Z,\kappa} \) can be represented as \( \{111, 112, \ldots, 1K1, 121, \ldots, 112, jkl, \ldots, JKL\} \), where \( l \in \{1, \ldots, L\} \) spans the state space of \( S^\kappa_t \). Let us denote this set of indexes \( \mathcal{R} \). The number of triplets in \( \mathcal{R} \) is \( Q \). For example, if \( J = 2 \), \( K = 3 \), and \( L = 4 \), the number of regimes in \( S_t^{\omega,Z,\kappa} \) is \( Q = 24 \), and the state space of \( S_t^{\omega,Z,\kappa} \) is represented by the indexes \( \mathcal{R} = \{111, 121, 131, 211, 221, 231, 112, 122, 132, 212, 222, 232, 113, 123, 133, 213, 223, 233, 114, 124, 134, 214, 224, 234\} \). Let us denote the generic element of \( \mathcal{R} \) as \( q_1q_2q_3 \), and its position as \( q = 1, 2, \ldots, Q \). For instance, when \( q_1 = 2 \), \( q_2 = 3 \), and \( q_3 = 1 \), the corresponding element in \( \mathcal{R} \) is \( \mathcal{R}_q = 231 \) with \( q = 6 \). The representation associated with \( S_t^{\omega,Z,\kappa} \) is reported in Figure 2. The path diagram displayed in Figure 2 is that of a standard HMM for which the conditional distribution of \( Y_{n,t} \) is given by a zero-inflated Skellam distribution. Specifically, conditional on \( S_t^{\omega,Z,\kappa} = \mathcal{R}_q \), we denote the probability mass function of \( Y_t \) as

\[
P(Y_t = y_t | S_t^{\omega,Z,\kappa} = \mathcal{R}_q) = \prod_{n=1}^{N} \left( \kappa_{n,t,q_3} \psi(y_{n,t}) + (1 - \kappa_{n,t,q_3}) SK(y_{n,t}, \lambda^{(1)}_{n,q_2}, \lambda^{(2)}_{n,q_2}) \right),
\]

which is adopted to make filtering and smoothing of the latent states as simple as it is for standard hidden Markov models. The mapping of the general DyMiSk model into the constrained representations is unique, given a suitable order in the state space of the Markov chains. In other words, all DyMiSk models admit the two simpler (restricted) representation outlined above. However, the reverse is not true: the simpler representations above do not necessarily map into a DyMiSk model representation. Note that the stationarity of the combined chain \( S_t^{\omega,Z,\kappa} \) follows from the stationarity of \( S_t^\omega \) and \( S_t^\kappa \).

### 2 Derivation of the EM algorithm

In order to develop an EM algorithm for the maximum likelihood (ML) estimation of the parameters of the DyMiSk model, we exploit the stochastic representation of the Skellam distribution as the difference between two independent Poisson random variables. Consider now the joint distribution of \( (Y_{n,t}, X_{n,t}^{(1)}) | (Z_t =
The EM algorithm treats these unobserved terms as missing values and proceeds with the estimation of

Unfortunately, this log-likelihood cannot be directly maximized due to the presence of latent quantities.

number of variables

matrix with the Bernoulli probabilities for all assets and regimes of

where

lam as the difference of two Poisson random variables,

vectors

By taking the logarithm and removing the quantities that do not depend on model parameters, we obtain

that is, the product of the two Poisson probability mass functions with intensity \( \lambda_{n:k}^{(1)} \) and \( \lambda_{n:k}^{(2)} \) evaluated in \( x_{n,t}^{(1)} \) and \( x_{n,t}^{(1)} - y_{n,t} \), respectively.

Let’s consider a sample of \( T \) observations for \( N \) price changes collected in the \( N \times T \) matrix \( y_{1:T} = (y_1, \ldots, y_T) \), and the series of random variables \( S^\omega_{1:T} = (S^\omega_1, \ldots, S^\omega_T)' \), \( S^\kappa_{1:T} = (S^\kappa_1, \ldots, S^\kappa_T)' \), \( Z_{1:T} = (Z_1, \ldots, Z_T)' \), and \( B_{1:T} = (B_1, \ldots, B_T) \), where their (unobserved) realizations are collected in the \( T \times 1 \) vectors \( s^\omega_{1:T}, s^\kappa_{1:T}, z_{1:T}, \) and in the \( N \times T \) matrix \( b_{1:T} \). To exploit the stochastic representation of the Skellam as the difference of two Poisson random variables,\(^1\) consider the \( N \times T \) matrix \( X_{1:T} = (X_{1:T}^{(1)}, \ldots, X_{1:T}^{(T)}) \), where \( X_{1:T}^{(t)} = (X_{1:T}^{(1)}, n = 1, \ldots, N)' \) and its (unobserved) realization \( x_{1:T}^{(t)} \). We collect all model parameters in the vector \( \theta = (\text{vec}(\Omega)', \text{vec}(\kappa)', \text{vec}(\Gamma^\omega)', \text{vec}(\Gamma^\kappa)', \text{vec}(\lambda^{(1)})', \text{vec}(\lambda^{(2)})') \), where \( \kappa = [\kappa_{n,t}] \) is a \( N \times L \) matrix with the Bernoulli probabilities for all assets and regimes of \( S^\kappa \). The number of free parameters is \( J(J - 1) \) for \( \Gamma^\omega \), \( L(L - 1) \) for \( \Gamma^\kappa \), \( J(K - 1) \) for \( \Omega \), \( LN \) for \( \kappa \) and \( 2KN \) for \( \lambda^{(1)} \) and \( \lambda^{(2)} \). This means that model complexity in terms of number of free parameters is quadratic in \( J \) and \( L \), while being linear in the number of variables \( N \) and mixture components \( K \).

The likelihood of observed and unobserved random variables, \( L(\theta|S^\omega_{1:T} = s^\omega_{1:T}, S^\kappa_{1:T} = s^\kappa_{1:T}, Z_{1:T} = z_{1:T}, B_{1:T} = b_{1:T}, X_{1:T}^{(1)} = x_{1:T}^{(1)}, Y_{1:T} = y_{1:T}) \), is

\[
L(\theta|\cdot) = \delta_{s^\omega_{1:T}}^{s^\omega_{1:T}} \left( \prod_{t=2}^{T} \gamma_{t-1}^{s^\omega_{t-1}, s^\omega_t} \right) \left( \prod_{t=2}^{T} \gamma_{t-1}^{s^\kappa_{t-1}, s^\kappa_t} \right) \prod_{t=1}^{T} \omega_{z_{t}, s^\kappa_t} \times \prod_{n=1}^{N} \psi(y_{n,t})^{b_{n,t}} \left( e^{-\left( \lambda_{n,z_{t}}^{(1)} + \lambda_{n,z_{t}}^{(2)} \right)} \frac{x_{n,t}^{(1)} \lambda_{n,z_{t}}^{(1)} (x_{n,t}^{(1)} - y_{n,t})}{x_{n,t}^{(1)}! (x_{n,t}^{(1)} - y_{n,t})!} \right)^{1-b_{n,t}} \kappa_{n,z_{t}}^{b_{n,t}} (1 - \kappa_{n,z_{t}})^{1-b_{n,t}},
\]

By taking the logarithm and removing the quantities that do not depend on model parameters, we obtain

\[
\log L(\theta|\cdot) \propto \log(\delta_{s^\omega_{1:T}}^{s^\omega_{1:T}}) + \log(\delta_{s^\kappa_{1:T}}^{s^\kappa_{1:T}}) + \sum_{t=1}^{T} \log(\omega_{z_{t}, s^\kappa_t}) + \sum_{t=2}^{T} \log(\gamma_{t-1}^{s^\kappa_{t-1}, s^\kappa_t}) \\
+ \sum_{t=2}^{T} \log(\gamma_{t-1}^{s^\omega_{t-1}, s^\omega_t}) + \sum_{n=1}^{N} \sum_{t=1}^{T} b_{n,t} \log(\kappa_{n,s^\kappa_t}) + \sum_{n=1}^{N} \sum_{t=1}^{T} (1 - b_{n,t}) \log(1 - \kappa_{n,s^\kappa_t}) \\
+ \sum_{n=1}^{N} \sum_{t=1}^{T} (1 - b_{n,t}) \left( -\left( \lambda_{n,z_{t}}^{(1)} + \lambda_{n,z_{t}}^{(2)} \right) + x_{n,t}^{(1)} \log(\lambda_{n,z_{t}}^{(1)}) + (x_{n,t}^{(1)} - y_{n,t}) \log(\lambda_{n,z_{t}}^{(2)}) \right).
\]

Unfortunately, this log-likelihood cannot be directly maximized due to the presence of latent quantities. The EM algorithm treats these unobserved terms as missing values and proceeds with the estimation of the expected value of the so-called complete data log-likelihood (CDLL). For the implementation of the EM

\(^1\)Lemma 1 in \(^7\) shows that the Skellam distribution does not necessarily arise from the difference of two independent Poisson random variables, but it can also be derived as the difference of random variables with a specific trivariate structure.
algorithm, we introduce the following additional variables
\[ u_{t,i,j}^\omega = \begin{cases} 1, & \text{if } S_{t}^\omega = i; \\ 0, & \text{otherwise.} \end{cases} \]
\[ u_{t,l}^\kappa = \begin{cases} 1, & \text{if } S_{t}^\kappa = l; \\ 0, & \text{otherwise.} \end{cases} \]
\[ z_{t,j,k} = \begin{cases} 1, & \text{if } Z_{t} = k, S_{t}^\omega = j; \\ 0, & \text{otherwise.} \end{cases} \]
\[ v_{t,i,j}^\omega = \begin{cases} 1, & \text{if } S_{t-1}^\omega = i, S_{t}^\omega = j; \\ 0, & \text{otherwise.} \end{cases} \]
\[ v_{t,l}^\kappa = \begin{cases} 1, & \text{if } S_{t-1}^\kappa = h, S_{t}^\kappa = l; \\ 0, & \text{otherwise.} \end{cases} \]

The variables \( u_{t,i,j}^\omega, u_{t,l}^\kappa, v_{t,i,j}^\omega \) and \( v_{t,l}^\kappa \) follow from the standard implementation of the algorithm for Markov-switching models, see ?, whereas the variable \( z_{t,j,k} \) (for \( j = 1, \ldots, J \) and \( k = 1, \ldots, K \)) is specific to our model and is related to the additional latent variable \( Z_{t} \). The new variables allow us to write the CDLL as

\[
\log L^C(\theta|\cdot) \propto \sum_{j=1}^{J} u_{t,j}^\omega \log(\delta_{j}^\omega) + \sum_{l=1}^{L} u_{t,l}^\kappa \log(\gamma_{l}) + \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{l=1}^{L} u_{t,n,l} \log(1 - b_{n,t,l}) \log(1 - \kappa_{n,l}) + \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{k=1}^{K} z_{t,j,k} \log(\omega_{j,k}) + \sum_{t=2}^{T} \sum_{j=1}^{J} \sum_{i=1}^{I} u_{t-1,i,j}^\omega \log(\delta_{j}^\omega) + \sum_{t=2}^{T} \sum_{l=1}^{L} \sum_{i=1}^{I} v_{t-1,i,j}^\omega \log(\gamma_{l,i}) + \sum_{t=2}^{T} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{k=1}^{K} u_{t-1,n,l,k} \log(\gamma_{n,l,k}) + \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{k=1}^{K} \sum_{j=1}^{J} \sum_{k=1}^{K} (1 - b_{n,t,l}) z_{t,j,k} \left( -\left( \lambda_{n,k}^{(1)} + \lambda_{n,k}^{(2)} \right) + x_{n,t,k}^{(1)} \log(\lambda_{n,k}^{(1)}) + \left( x_{n,t,k}^{(1)} - y_{n,t} \right) \log(\lambda_{n,k}^{(2)}) \right) \]

where \( b_{n,t,l} \) and \( x_{n,t,k}^{(1)} \) indicate the realizations of \( B_{n,t}|S_{t}^\omega = l \) and of \( X_{n,t,k}^{(1)} | Z_{t} = k \), respectively.

The EM algorithm iterates between the expectation-step (E-step) and maximization-step (M-Step) until convergence. Given a value of the model parameters at iteration \( m \), \( \Theta^{(m)} \), the E-step consists of the evaluation of the so-called Q function defined as \( Q(\theta, \theta^{(m)}) = E_{\theta^{(m)}}(\log L^C(\theta|\cdot)) \), where the expectation is taken with respect to the joint distribution of the missing variables conditional to the observed variables using parameter values at iteration \( m \). Exploiting the formulation of the CDLL in (3), the Q function can be factorized as

\[
Q(\theta, \theta^{(m)}) \propto \sum_{j=1}^{J} \tilde{u}_{1,j}^\omega \log(\delta_{j}^\omega) + \sum_{l=1}^{L} \tilde{u}_{1,l}^\kappa \log(\gamma_{l}) + \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{l=1}^{L} \tilde{u}_{t,n,l} \log(1 - \tilde{b}_{n,t,l}) \log(1 - \kappa_{n,l}) + \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{k=1}^{K} \tilde{z}_{t,j,k} \log(\omega_{j,k}) + \sum_{t=2}^{T} \sum_{j=1}^{J} \sum_{i=1}^{I} \tilde{u}_{t-1,i,j}^\omega \log(\delta_{j}^\omega) + \sum_{t=2}^{T} \sum_{l=1}^{L} \sum_{i=1}^{I} \tilde{v}_{t-1,i,j}^\omega \log(\gamma_{l,i}) + \sum_{t=2}^{T} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{k=1}^{K} \tilde{u}_{t-1,n,l,k} \log(\gamma_{n,l,k}) + \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{k=1}^{K} \sum_{j=1}^{J} \sum_{k=1}^{K} (1 - \tilde{b}_{n,t,l}) \tilde{z}_{t,j,k} \left( -\left( \tilde{\lambda}_{n,k}^{(1)} + \tilde{\lambda}_{n,k}^{(2)} \right) + \tilde{x}_{n,t,k}^{(1)} \log(\tilde{\lambda}_{n,k}^{(1)}) + \left( \tilde{x}_{n,t,k}^{(1)} - y_{n,t} \right) \log(\tilde{\lambda}_{n,k}^{(2)}) \right) \]

where \( \tilde{u}_{t,j}^\omega = P(S_{t}^\omega = j|Y_{1:T} = Y_{1:T}), \tilde{u}_{t,l}^\kappa = P(S_{t}^\kappa = l|Y_{1:T} = Y_{1:T}), \tilde{v}_{t,i,j}^\omega = P(S_{t-1}^\omega = i, S_{t}^\omega = j|Y_{1:T} = Y_{1:T}), \tilde{v}_{t,l}^\kappa = P(S_{t-1}^\kappa = h, S_{t}^\kappa = l|Y_{1:T} = Y_{1:T}), \tilde{z}_{t,j,k} = P(Z_{t} = k, S_{t}^\omega = j|Y_{1:T} = Y_{1:T}), \tilde{b}_{n,t,l} = P(B_{n,t,l} = 1|S_{t}^\omega = l, Y_{1:T} = Y_{1:T}), \text{and } \tilde{x}_{n,t,k}^{(1)} = E[X_{n,t,k}^{(1)} | Z_{t} = k, Y_{1:T} = Y_{1:T}]. \) The E-step involves the computation of these quantities.

Furthermore, let us define the forward and backward probabilities for the third model representation reported in Figure 2 and can be evaluated using the FFBS algorithm. In particular, the forward probabilities
are $\alpha_{t,q} = P(S_t^q, Z_t, \kappa = R_q, Y_{1:t} = y_{1:t})$ with initial condition $\alpha_{0,q} = \delta_{q,R_q}$ for all $q = 1, \ldots, JKL$, where $\delta_{q,R_q} = P(S_0^q, Z_0, \kappa = R_q)$ is the stationary distribution of $S_t^q, Z_t, \kappa$ in state $R_q$. The backward probabilities are $\beta_{T,q} = P(Y_{t+1:T} = y_{t+1:T} | S_t^q, Z_t, \kappa = R_q)$ with final condition $\beta_{T,q} = 1$. Once these probabilities are evaluated, the following smoothed probabilities can be computed

$$P(S_t^q, Z_t, \kappa = R_q | Y_{1:T} = y_{1:T}) = \frac{\alpha_{t,q} \beta_{T,q}}{\sum_{r=1}^{JKL} \alpha_{t,r} \beta_{T,r}},$$

$$P(S_t^q, Z_t, \kappa = R_s | Y_{1:T} = y_{1:T}) = \frac{\alpha_{t-1:s} \gamma_{j,s}^q P(Y_t = y_t | S_t^q, Z_t, \kappa = R_q) \beta_{T,q}}{\sum_{r=1}^{JKL} \alpha_{t,r} \beta_{T,r}},$$

where $\gamma_{j,s}^q$ is the $(j,s)$-th element of $\Gamma_t^q$, and $P(Y_t = y_t | S_t^q, Z_t, \kappa = R_q)$ is given in equation (2). Since $P(S_t^q, Z_t, \kappa = R_q | Y_{1:T}) = P(S_t^q, Z_t, \kappa = R_{1:T}, Y_{1:T})$ and $P(S_t^q, Z_t, \kappa = R_q | Y_{1:T}) = P(S_t^q, Z_t, \kappa = R_{1:T}, Y_{1:T})$ it follows that $\tilde{w}_{t,i,j}$, $\tilde{w}_{t,h,l}$, and $\tilde{\gamma}_{t,i,j}$ can be evaluated by simple marginalization of the relevant variables from $P(S_t^q, Z_t, \kappa = R_q | Y_{1:T})$ and $P(S_t^q, Z_t, \kappa = R_q | Y_{1:T})$. It also follows that the joint probabilities $P(S_t^q = j, Z_t = k, S_t^q = l | Y_{1:T})$ are immediately available. The remaining quantities are given by

$$\tilde{b}_{n,t,l} = \begin{cases} 0, & \text{if } y_{n,t} \neq 0 \\ \sum_{k=1}^{K} \sum_{j=1}^{J} \frac{\kappa_{n,j} P(S_t^q = j, Z_t = k, S_t^q = l | Y_{1:T})}{P(S_t^q = l | Y_{1:T})(\kappa_{n,j} \psi(y_n) + (1 - \kappa_{n,j}) SK(y_n, \gamma_{n,k}^{(1)}(y_n)^{(2)}))}, & \text{otherwise}, \end{cases}$$

and

$$\tilde{x}_{n,t,k}^{(1)} = \sum_{j=1}^{J} \sum_{l=1}^{L} \lambda_{n,k} SK(y_n - 1, \gamma_{n,k}^{(1)}(y_n), \gamma_{n,k}^{(2)}(y_n)) P(S_t^q = j, Z_t = k, S_t^q = l | Y_{1:T} = y_{1:t}).$$

In the M-step of the algorithm, the function $Q$ is maximized with respect to the model parameters $\theta$. Solving the Lagrangian associated with this (constrained) optimization leads to the following solution of the maximization problem:

$$\gamma_{i,j}^{(m+1)} = \frac{\sum_{t=1}^{T} \tilde{w}_{t,i,j}^{(m)}}{\sum_{t=1}^{T} \tilde{w}_{t,i,j}}, \quad \gamma_{h,l}^{(m+1)} = \frac{\sum_{t=1}^{T} \tilde{w}_{t,h,l}^{(m)}}{\sum_{t=1}^{T} \tilde{w}_{t,h,l}}, \quad \kappa_{n,j}^{(m+1)} = \frac{\sum_{t=1}^{T} \tilde{b}_{n,t,l}^{(m)}}{\sum_{t=1}^{T} \tilde{u}_{n,t}^{(m)}},$$

$$\delta_{j}^{(m+1)} = \frac{\sum_{t=1}^{T} \tilde{z}_{t,j,k}^{(1)}(1 - \tilde{b}_{n,t,l}) y_{n,t}}{\sum_{t=1}^{T} \tilde{z}_{t,j,k}^{(1)}(1 - \tilde{b}_{n,t,l})}, \quad \delta_{j}^{\omega,(m+1)} = \tilde{u}_{t,j}^{(m+1)}.$$
3 Estimated parameters

In the following the \( \hat{\lambda}_i = [\hat{\lambda}_{n,k}] \) for \( n = 1, \ldots, 4, k = 1, \ldots, K \) and \( i = 1, 2 \), \( \hat{\Omega} = [\hat{\omega}_{k,j}] \) for \( k = 1, \ldots, K \) and \( j = 1, \ldots, J \), \( \hat{\kappa}_l = [\hat{\kappa}_{d,n}] \) for \( l = 1, \ldots, L, d = 1, \ldots, D \), and \( n = 1, \ldots, N \), \( \hat{\beta} = [\hat{\beta}_{d,n}] \) for \( d = 1, \ldots, D \), with \( D = 14 \) in all cases, and \( n = 1, \ldots, N \) with \( N = 4 \).

Lehman period

The model selected by BIC has \( J = 12, L = 2, K = 5 \). The estimated coefficients are:

\[
\begin{align*}
\hat{\lambda}_1 &= \begin{pmatrix}
1.474 & 1.301 & 4.681 & 3.292 \\
1.242 & 1.126 & 10.557 & 6.034 \\
2.063 & 1.992 & 27.940 & 12.727 \\
3.418 & 3.092 & 11.107 & 8.415 \\
4.431 & 9.773 & 28.873 & 12.663 \\
10.026 & 3.126 & 24.540 & 17.268 \\
7.264 & 8.022 & 36.079 & 23.348 \\
5.037 & 3.117 & 29.575 & 17.162 \\
7.857 & 8.382 & 44.064 & 18.010 \\
17.156 & 18.560 & 63.703 & 41.231 \\
52.323 & 53.428 & 113.974 & 97.833 \\
31.954 & 50.414 & 108.937 & 76.746
\end{pmatrix}, \\
\hat{\lambda}_2 &= \begin{pmatrix}
1.607 & 1.346 & 4.756 & 3.453 \\
3.597 & 3.218 & 14.151 & 9.933 \\
1.764 & 1.686 & 27.524 & 11.738 \\
0.764 & 0.950 & 7.529 & 4.442 \\
3.320 & 6.438 & 28.328 & 11.066 \\
2.185 & 1.252 & 22.047 & 16.597 \\
1.838 & 2.908 & 27.355 & 13.667 \\
10.423 & 8.577 & 39.013 & 27.512 \\
8.919 & 10.076 & 43.602 & 19.281 \\
16.867 & 19.260 & 86.463 & 43.607 \\
57.907 & 65.418 & 115.000 & 111.262 \\
28.891 & 44.999 & 102.079 & 68.101
\end{pmatrix}, \\
\hat{\Omega} &= \begin{pmatrix}
0.590 & 0.147 & 0.137 & 0.025 & 0.000 \\
0.190 & 0.062 & 0.242 & 0.156 & 0.000 \\
0.002 & 0.095 & 0.396 & 0.119 & 0.001 \\
0.185 & 0.008 & 0.135 & 0.062 & 0.000 \\
0.001 & 0.014 & 0.002 & 0.141 & 0.079 \\
0.000 & 0.036 & 0.000 & 0.006 & 0.013 \\
0.000 & 0.010 & 0.048 & 0.157 & 0.030 \\
0.008 & 0.008 & 0.038 & 0.126 & 0.012 \\
0.000 & 0.047 & 0.002 & 0.204 & 0.373 \\
0.000 & 0.001 & 0.000 & 0.005 & 0.080 \\
0.000 & 0.003 & 0.000 & 0.000 & 0.128 \\
0.000 & 0.028 & 0.000 & 0.000 & 0.285
\end{pmatrix}, \\
\hat{\beta} &= \begin{pmatrix}
1.539 & 1.473 & 1.575 & 1.920 \\
1.070 & 0.830 & 1.061 & 1.300 \\
1.117 & 0.807 & 1.037 & 1.108 \\
0.991 & 0.675 & 0.951 & 0.978 \\
0.796 & 0.649 & 0.974 & 0.864 \\
0.925 & 0.612 & 0.718 & 0.761 \\
0.661 & 0.443 & 0.630 & 0.611 \\
0.693 & 0.515 & 0.629 & 0.587 \\
0.741 & 0.494 & 0.712 & 0.655 \\
0.829 & 0.643 & 0.981 & 0.861 \\
0.802 & 0.647 & 1.027 & 0.812 \\
0.862 & 0.586 & 0.956 & 0.874 \\
0.969 & 0.756 & 1.281 & 0.927
\end{pmatrix}
\end{align*}

The model selected by BIC has $J = 5$, $L = 1$ ($\Gamma^k = 1$), and $K = 5$. The estimated coefficients are:

$$
\hat{\kappa}_1 = \begin{pmatrix}
0.090 & 0.894 & 0.842 & 0.975 \\
0.049 & 0.867 & 0.596 & 1.000 \\
0.089 & 0.161 & 0.074 & 0.034 \\
0.071 & 0.114 & 0.000 & 0.040 \\
0.081 & 0.175 & 0.024 & 0.043 \\
0.043 & 0.166 & 0.000 & 0.029 \\
0.105 & 0.147 & 0.076 & 0.026 \\
0.091 & 0.002 & 0.145 & 0.140 \\
0.065 & 0.039 & 0.069 & 0.082 \\
0.112 & 0.080 & 0.175 & 0.024 \\
0.028 & 0.158 & 0.026 & 0.100 \\
0.007 & 0.172 & 0.000 & 0.059 \\
0.030 & 0.135 & 0.000 & 0.046 \\
0.009 & 0.054 & 0.058 & 0.000
\end{pmatrix}, \quad \hat{\kappa}_2 = \begin{pmatrix}
0.031 & 0.093 & 0.016 & 0.114 \\
0.032 & 0.030 & 0.012 & 0.033 \\
0.001 & 0.025 & 0.006 & 0.033 \\
0.012 & 0.000 & 0.022 & 0.034 \\
0.006 & 0.000 & 0.021 & 0.022 \\
0.010 & 0.000 & 0.032 & 0.027 \\
0.023 & 0.001 & 0.010 & 0.077 \\
0.014 & 0.000 & 0.012 & 0.029 \\
0.000 & 0.037 & 0.019 & 0.015 \\
0.020 & 0.008 & 0.015 & 0.042 \\
0.000 & 0.000 & 0.025 & 0.001 \\
0.047 & 0.000 & 0.078 & 0.029 \\
0.000 & 0.000 & 0.037 & 0.019 \\
0.000 & 0.000 & 0.028 & 0.023
\end{pmatrix}
$$

Low Volatility period

The model selected by BIC has $J = 5$, $L = 1$ ($\Gamma^k = 1$), and $K = 5$. The estimated coefficients are:

$$
\hat{\Gamma}^k = \begin{pmatrix}
0.984 & 0.004 & 0.008 & 0.000 & 0.004 \\
0.000 & 0.979 & 0.001 & 0.001 & 0.019 \\
0.035 & 0.004 & 0.957 & 0.004 & 0.000 \\
0.000 & 0.002 & 0.003 & 0.995 & 0.001 \\
0.000 & 0.012 & 0.001 & 0.001 & 0.987
\end{pmatrix}, \quad \hat{\Omega} = \begin{pmatrix}
0.532 & 0.302 & 0.061 & 0.102 & 0.782 \\
0.161 & 0.408 & 0.038 & 0.308 & 0.192 \\
0.053 & 0.255 & 0.056 & 0.436 & 0.022 \\
0.002 & 0.034 & 0.035 & 0.150 & 0.005 \\
0.252 & 0.001 & 0.809 & 0.004 & 0.000
\end{pmatrix}
$$

$$
\hat{\lambda}_1 = \begin{pmatrix}
0.651 & 0.522 & 0.650 & 0.901 \\
0.965 & 0.803 & 1.347 & 2.078 \\
0.632 & 0.428 & 0.723 & 1.328 \\
2.345 & 1.530 & 3.122 & 4.649 \\
5.421 & 0.928 & 2.450 & 3.117
\end{pmatrix}, \quad \hat{\lambda}_2 = \begin{pmatrix}
0.755 & 0.597 & 0.796 & 1.148 \\
0.425 & 0.407 & 0.620 & 0.907 \\
1.578 & 1.074 & 2.012 & 3.187 \\
0.392 & 0.156 & 0.227 & 1.921 \\
5.284 & 0.992 & 2.377 & 3.107
\end{pmatrix}
$$

$$
\hat{\kappa}_1 = \begin{pmatrix}
0.063 & 0.000 & 0.032 & 0.037 \\
0.010 & 0.007 & 0.035 & 0.034 \\
0.031 & 0.000 & 0.040 & 0.049 \\
0.003 & 0.000 & 0.000 & 0.026 \\
0.036 & 0.000 & 0.001 & 0.052 \\
0.063 & 0.000 & 0.000 & 0.002 \\
0.001 & 0.000 & 0.002 & 0.076 \\
0.003 & 0.000 & 0.000 & 0.002 \\
0.001 & 0.000 & 0.000 & 0.040 \\
0.000 & 0.000 & 0.000 & 0.001 \\
0.004 & 0.000 & 0.000 & 0.006 \\
0.019 & 0.000 & 0.000 & 0.051 \\
0.000 & 0.000 & 0.000 & 0.000 \\
0.004 & 0.000 & 0.000 & 0.000
\end{pmatrix}, \quad \hat{\beta} = \begin{pmatrix}
1.740 & 1.346 & 2.059 & 2.581 \\
1.393 & 0.820 & 1.234 & 1.381 \\
1.062 & 0.622 & 0.738 & 0.950 \\
0.984 & 0.675 & 0.696 & 0.853 \\
0.878 & 0.563 & 0.692 & 0.726 \\
0.638 & 0.594 & 0.597 & 0.707 \\
0.525 & 0.507 & 0.535 & 0.591 \\
0.528 & 0.525 & 0.579 & 0.523 \\
0.542 & 0.517 & 0.477 & 0.479 \\
0.527 & 0.521 & 0.468 & 0.462 \\
0.573 & 0.534 & 0.541 & 0.477 \\
0.552 & 0.586 & 0.583 & 0.439 \\
0.666 & 0.600 & 0.598 & 0.478
\end{pmatrix}
$$
4 Additional figures and tables

4.1 Univariate distributions

Figure 3: Comparison between the empirical and model-implied unconditional distribution of CAT for the low volatility period. The first figure reports the distribution of the first 5 minutes of trading activity (9:30 - 9:35), the second reports the distribution for the following 25 minutes (9:35 - 10:00), figures from the third to the fifteen display the distribution computed every 30 minutes. Yellow lines represent the probability implied by the unconditional distribution of the DyMiSk model.

(a) Low volatility - In-Sample

(b) Low volatility - Out of Sample

Figure 4: PITs. In-sample and out-of-sample randomized PITs as in ? computed according to the one step ahead univariate conditional distribution of each asset. PITs are divided in 10 bins such that under the null hypothesis of correct model specification the area of each bin should be 10%. Confidence intervals based on the methodology of ? are computed at the 5% level.
Figure 5: Comparison between the empirical and model-implied unconditional distribution of KO the Low Volatility (panel (a)) and the Lehman (panel (b)) period. The first figure of each panel reports the distribution of the first 5 minutes of trading activity (9:30 - 9:35), the second reports the distribution for the following 25 minutes (9:35 - 10:00), figures from the third to the fifteenth display the distribution computed every 30 minutes. Yellow lines represent the probability implied by the unconditional distribution of the DMS model.
Figure 6: Comparison between the empirical and model implied unconditional distribution of JPM the Low Volatility (panel (a)) and the Lehman (panel (b)) period. The first figure of each panel reports the distribution of the first 5 minutes of trading activity (9:30 - 9:35), the second reports the distribution for the following 25 minutes (9:35 - 10:00), figures from the third to the fifteen display the distribution computed every 30 minutes. Yellow lines represent the probability implied by the unconditional distribution of the DMS model.
Figure 7: Comparison between the empirical and model implied unconditional distribution of WMT the Low Volatility (panel (a)) and the Lehman (panel (b)) period. The first figure of each panel reports the distribution of the first 5 minutes of trading activity (9:30 - 9:35), the second reports the distribution for the following 25 minutes (9:35 - 10:00), figures from the third to the fifteen display the distribution computed every 30 minutes. Yellow lines represent the probability implied by the unconditional distribution of the DMS model.
4.2 Bivariate Distributions

Figure 8: Bivariate unconditional distributions. The Figure reports the empirical and model-based unconditional distributions of KO and JPM during the opening of the market (9:30-09:35), the lunch time (12:00-12:30), and the closing (15:30-16:00) for both the Low Volatility (Panels a-c) and the Lehman (Panels d-f) period. The full red circles represent the empirical frequencies computed over the estimation period. The blue circles represent the theoretical frequencies computed according to the unconditional bivariate distribution of KO and JPM.

Figure 9: Bivariate unconditional distributions. The Figure reports the empirical and model-based unconditional distributions of WMT and JPM during the opening of the market (9:30-09:35), the lunch time (12:00-12:30), and the closing (15:30-16:00) for both the Low Volatility (Panels a-c) and the Lehman (Panels d-f) period. The full red circles represent the empirical frequencies computed over the estimation period. The blue circles represent the theoretical frequencies computed according to the unconditional bivariate distribution of WMT and JPM.
4.3 Volatility and Correlations

Figure 10: Absolute price changes (red squares) with in-sample predicted volatility, \( \hat{\sigma}_{t|t-1,i} \), (blue solid line) during the tranquil period. Predicted volatilities are computed according to the one step ahead conditional distribution over the in-sample period.

Figure 11: Absolute price changes (red squares) with out-of-sample predicted volatility, \( \sigma_{n,t|t-1} \), (blue solid line) during the Lehman period.
Figure 12: Filtered correlation after aggregation (blue dashed line) and realized correlation (black solid line) in the Lehman (top) and the Low Volatility (bottom) periods of WMT versus other assets. The correlation measures are constructed aggregating $\hat{\Sigma}_{t|t-1}$ over 30-minutes intervals. The realized correlations are computed starting from the realized covariance matrix estimated with the realized kernel estimator of Barndorff-Nielsen et al. (2008) based on price changes.

Figure 13: In-sample randomized PITs as in ?.
| Asset | (a)        | (b)        | (c)        | (d)        |
|-------|------------|------------|------------|------------|
|       | Intercept  | Intercept  | Intercept  | Intercept  |
| AXP   | -2.28***   | -2.54***   | -2.38***   | -2.27***   |
|       | \(\hat{P}_{t-1}\) | 4.54***   | 5.24***   | 4.38***   |
|       | Dummy      | ✓          | ✓          | ✓          |
|       | Lags       | ✓          | ✓          | ✓          |
|       | Bid Ask    | ✓          | ✓          | ✓          |
|       | Pseudo R\(^2\) | 0.02      | 0.02      | 0.02      |
|       |            |            |            |            |
| UTX   | -2.65***   | -2.08***   | -2.00***   | -1.48***   |
|       | \(\hat{P}_{t-1}\) | 6.03***   | 3.46**    | 2.73*     |
|       | Dummy      | ✓          | ✓          | ✓          |
|       | Lags       | ✓          | ✓          | ✓          |
|       | Bid Ask    | ✓          | ✓          | ✓          |
|       | Pseudo R\(^2\) | 0.01      | 0.01      | 0.01      |
| PFE   | -1.86***   | -2.16***   | -2.00***   | -2.28***   |
|       | \(\hat{P}_{t-1}\) | 3.53***   | 4.14***   | 3.46***   |
|       | Dummy      | ✓          | ✓          | ✓          |
|       | Lags       | ✓          | ✓          | ✓          |
|       | Bid Ask    | ✓          | ✓          | ✓          |
|       | Pseudo R\(^2\) | 0.01      | 0.01      | 0.01      |
| GS    | -4.32***   | -3.87***   | -3.84***   | -2.95***   |
|       | \(\hat{P}_{t-1}\) | 26.73***  | 16.01     | 15.10     |
|       | Dummy      | ✓          | ✓          | ✓          |
|       | Lags       | ✓          | ✓          | ✓          |
|       | Bid Ask    | ✓          | ✓          | ✓          |
|       | Pseudo R\(^2\) | 0.01      | 0.01      | 0.01      |
| BA    | -2.38***   | -2.47***   | -2.16***   | -1.53***   |
|       | \(\hat{P}_{t-1}\) | 4.96***   | 5.06***   | 2.92***   |
|       | Dummy      | ✓          | ✓          | ✓          |
|       | Lags       | ✓          | ✓          | ✓          |
|       | Bid Ask    | ✓          | ✓          | ✓          |
|       | Pseudo R\(^2\) | 0.01      | 0.01      | 0.01      |
| IBM   | -3.46***   | -3.85***   | -3.69***   | -2.59***   |
|       | \(\hat{P}_{t-1}\) | 12.49***  | 16.77***  | 14.05***  |
|       | Dummy      | ✓          | ✓          | ✓          |
|       | Lags       | ✓          | ✓          | ✓          |
|       | Bid Ask    | ✓          | ✓          | ✓          |
|       | Pseudo R\(^2\) | 0.01      | 0.01      | 0.01      |
| TRV   | -3.24***   | -3.55***   | -3.36***   | -2.97***   |
|       | \(\hat{P}_{t-1}\) | 10.09***  | 12.05***  | 10.03***  |
|       | Dummy      | ✓          | ✓          | ✓          |
|       | Lags       | ✓          | ✓          | ✓          |
|       | Bid Ask    | ✓          | ✓          | ✓          |
|       | Pseudo R\(^2\) | 0.01      | 0.01      | 0.01      |
| AAPL  | -3.76***   | -3.99***   | -3.90***   | -3.14***   |
|       | \(\hat{P}_{t-1}\) | 16.21***  | 19.68***  | 17.82***  |
|       | Dummy      | ✓          | ✓          | ✓          |
|       | Lags       | ✓          | ✓          | ✓          |
|       | Bid Ask    | ✓          | ✓          | ✓          |
|       | Pseudo R\(^2\) | 0.01      | 0.01      | 0.01      |

Table 1: Estimated coefficients of the logistic regression in equation (11). The Table reports the results for each asset considered in Section 4.4. We consider the regression in equation (11) with no control variables (a), with seasonal dummies (b), with seasonal dummies and 5 lags of the dependent variable (c), with seasonal dummies, lags of the dependent variable and the bid-ask spread (d). The superscript *** *, **, and * indicate statistical significance at the 1%, 5%, and 10% significance levels, respectively. The standard errors are computed according to the Newey-West formula (HAC). The pseudo R\(^2\) is the goodness-of-fit index by ?.
References

Brockwell, A. (2007). Universal residuals: A multivariate transformation. *Statistics & Probability Letters, 77*(14):1473–1478.

Diebold, F. X., Gunther, T. A., and Tay, A. S. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review, 39*:863–883.

Karlis, D. and Ntzoufras, I. (2008). Bayesian modelling of football outcomes: using the Skellam’s distribution for the goal difference. *IMA Journal of Management Mathematics, 20*(2):133–145.

McFadden, D. (1974). *Analysis of Qualitative Choice Behavior*. Frontiers in Econometrics, Academic Press. New York, NY.

McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. John Wiley & Sons.