

# **Analyse bayésienne d'un modèle pour les changements de prix de titres échangés à haute fréquence sur les marchés financiers**

Raphaël Gendron  
raphael.gendron@umontreal.ca

Dirigé par William J. McCausland

Département des Sciences Économiques

Faculté des Arts et Sciences

Université de Montréal

Décembre 2005

# Sommaire

Cette recherche traite de l'analyse bayésienne du modèle de décomposition des changements de prix ADS pour des titres échangés à haute fréquence sur les marchés financiers. Ce modèle fut développé par Rydberg et Shephard (1998) et a déjà été estimé par la méthode classique du maximum de vraisemblance. La décomposition des changements de prix se fait en trois composantes (ADS) : présence d'un changement de prix (activity), direction du changement et ampleur (size). Afin de modéliser chacune de ces composantes, des variables de retard et de la microstructure des marchés sont utilisées.

Les paramètres des modèles sont estimés selon l'approche bayésienne pour les prix de l'action d'IBM transigée entre le 1<sup>er</sup> novembre 1990 et le 31 janvier 1991. Les données sont prises du New York Stock Exchange (NYSE) et on compte 59 838 observations pour cette période. L'outil de simulation Markov Chain Monte Carlo (MCMC) utilisé est l'algorithme de la marche aléatoire Metropolis-Hastings (M-H). La loi des tirages de cet algorithme converge en distribution vers la loi à posteriori.

# Table des matières

Sommaire	ii
Liste des tableaux	iv
Liste des graphiques	v
Liste des graphiques	v
1 Introduction	1
2 Informations pertinentes	3
2.1 Données	3
2.2 Éléments de l'analyse bayésienne	3
2.3 Algorithme de Metropolis-Hastings	4
3 Revue de la littérature	7
3.1 Le modèle probit ordonné	7
3.2 Le modèle de décomposition en trois composantes ADS	8
3.3 Études bayésiennes	11
4 Analyse théorique	12
4.1 La vraisemblance du modèle ADS	12
4.2 Le choix de la loi à priori	14
5 Analyse empirique	15
5.1 Ajustement des données	15
5.2 Résultats des estimations par MLE	16
5.3 Résultats des estimations par l'analyse bayésienne	18
5.3.1 Résultats pour la composante A	20
5.3.2 Résultats pour la composante D	27
5.3.3 Résultats pour la composante S	34
6 Conclusion	42
Bibliographie	44

## Liste des tableaux

Tableau 1: Estimation des paramètres du modèle ADS par maximum de vraisemblance pour l'action IBM échangée entre le 01/11/90 et le 31/01/91	17
Tableau 2: Estimation bayésienne des paramètres du modèle (5.11) pour les 25 000 premières transactions de l'action IBM	20
Tableau 3: Estimation bayésienne des paramètres du modèle (5.12) pour les 25 000 premières transactions de l'action IBM	21
Tableau 4: Estimation bayésienne des paramètres du modèle (5.13) pour les 25 000 premières transactions de l'action IBM	21
Tableau 5: Estimation bayésienne des paramètres du modèle (5.14) pour les 25 000 premières transactions de l'action IBM	21
Tableau 6 : Estimation bayésienne des paramètres du modèle (5.15) pour les 8 019 premières transactions avec changement de prix de l'action IBM	27
Tableau 7: Estimation bayésienne des paramètres du modèle (5.16) pour les 8 019 premières transactions avec changement de prix de l'action IBM	28
Tableau 8 : Estimation bayésienne des paramètres du modèle (5.17) pour les 8 019 premières transactions avec changement de prix de l'action IBM	28
Tableau 9 : Estimation bayésienne des paramètres du modèle (5.18) pour les 8 019 premières transactions avec changement de prix de l'action IBM	28
Tableau 10 : Estimation bayésienne des paramètres du modèle (5.19) pour les 4 023 premières transactions avec hausse de prix de l'action IBM	34
Tableau 11 : Estimation bayésienne des paramètres du modèle (5.20) pour les 4 023 premières transactions avec hausse de prix de l'action IBM	35
Tableau 12 : Estimation bayésienne des paramètres du modèle (5.21) pour les 4 023 premières transactions avec hausse de prix de l'action IBM	35
Tableau 13 : Estimation bayésienne des paramètres du modèle (5.22) pour les 4 023 premières transactions avec hausse de prix de l'action IBM	35
Tableau 14 : Estimation bayésienne des paramètres du modèle (5.19) pour les 3 996 premières transactions avec baisse de prix de l'action IBM	36
Tableau 15 : Estimation bayésienne des paramètres du modèle (5.20) pour les 3 996 premières transactions avec baisse de prix de l'action IBM	36
Tableau 16 : Estimation bayésienne des paramètres du modèle (5.21) pour les 3 996 premières transactions avec baisse de prix de l'action IBM	36
Tableau 17 : Estimation bayésienne des paramètres du modèle (5.22) pour les 3 996 premières transactions avec baisse de prix de l'action IBM	37

## Liste des graphiques

Graphique 1 : Distribution à posteriori du paramètre $\beta_0$ dans le modèle (5.12) pour les 25 000 premières transactions de l'action IBM _____	23
Graphique 2 : Distribution à posteriori du paramètre $\beta_1$ dans le modèle (5.12) pour les 25 000 premières transactions de l'action IBM _____	24
Graphique 3 : Trace du paramètre $\beta_0$ dans le modèle (5.12) lors de l'algorithme M-H pour les 25 000 premières transactions de l'action IBM _____	25
Graphique 4 : Trace du paramètre $\beta_1$ dans le modèle (5.12) lors de l'algorithme M-H pour les 25 000 premières transactions de l'action IBM _____	26
Graphique 5 : Distribution à posteriori du paramètre $\gamma_0$ dans le modèle (5.16) pour les 8 019 premières transactions avec changement de prix de l'action IBM _____	30
Graphique 6 : Distribution à posteriori du paramètre $\gamma_1$ dans le modèle (5.16) pour les 8 019 premières transactions avec changement de prix de l'action IBM _____	31
Graphique 7 : Trace du paramètre $\gamma_0$ dans le modèle (5.16) lors des 400 derniers tirages de l'algorithme M-H pour les 8 019 premières transactions avec changement de prix de l'action IBM _____	32
Graphique 8 : Trace du paramètre $\gamma_1$ dans le modèle (5.16) lors des 400 derniers tirages de l'algorithme M-H pour les 8 019 premières transactions avec changement de prix de l'action IBM _____	33
Graphique 9 : Distribution à posteriori du paramètre $\theta_{u,2}$ dans le modèle (5.21) pour les 4 023 premières transactions avec hausse de prix de l'action IBM _____	38
Graphique 10 : Distribution à posteriori du paramètre $\theta_{d,2}$ dans le modèle (5.21) pour les 3 996 premières transactions avec baisse de prix de l'action IBM _____	39
Graphique 11 : Trace du paramètre $\theta_{u,2}$ dans le modèle (5.21) lors de l'algorithme M-H pour les 4 023 premières transactions avec hausse de prix de l'action IBM _____	40
Graphique 12 : Trace du paramètre $\theta_{d,2}$ dans le modèle (5.21) lors de l'algorithme M-H pour les 3 996 premières transactions avec baisse de prix de l'action IBM _____	41

# 1 Introduction

Les données financières de haute fréquence présentent des caractéristiques importantes du fait qu'elles peuvent être observées à des intervalles de temps très précis. On les retrouve principalement dans le marché des titres financiers car chaque échange de ces titres est enregistré dans une base de données. Ces données sont également très importantes dans l'étude empirique de la microstructure des marchés financiers. Il est donc primordial d'étudier et de développer des modèles avec ce type de données car leurs caractéristiques uniques disparaissent lorsqu'elles sont agrégées dans le temps. Tsay (2002, p.183) remarque les principales caractéristiques suivantes :

- Les transactions sont enregistrées dans des intervalles de temps différents. Dans les données de basse fréquence, on retrouve généralement un prix aux 24 heures suite à l'agrégation des transactions;
- Le prix des actifs échangés est une valeur discrète et les changements du prix sont mesurés en fractions d'un dollar;
- La présence d'un cycle dans la quantité de titres échangés au cours d'une journée. En général, on remarque un grand nombre de transactions à l'ouverture et à la fermeture des heures régulières de transactions et un nombre plus faible au milieu de la journée;
- La possibilité qu'il y ait plusieurs transactions au cours d'une même seconde. Le prix d'un actif peut alors changer durant cette même seconde.

Dans le cadre de ce rapport de recherche, la caractéristique étudiée est le changement de prix des titres échangés à haute fréquence. Les données choisies sont les transactions de l'action d'IBM enregistrées au New York Stock Exchange (NYSE) du 1<sup>er</sup> novembre 1990 au 31 janvier 1991. On y dénombre 63 jours de transactions et 59 838 échanges de l'action. Malgré le grand nombre d'observations, le peu de véritables connaissances à priori concernant le comportement des prix des actifs échangés à haute fréquence

augmente le niveau de difficulté pour choisir le modèle le plus adéquat. Cette situation motive le choix d'une analyse bayésienne. Ce type d'analyse permet d'obtenir des résultats exacts pour des échantillons finis à l'aide de simulations et nous évite d'utiliser des lois asymptotiques. De plus, l'estimation bayésienne est dans un cadre probabiliste et tient automatiquement compte de l'incertitude des paramètres dans les prévisions. Le développement d'outils informatiques puissants contribue à la popularité croissante de cette méthode d'estimation. L'outil de simulation Markov Chain Monte Carlo (MCMC) utilisé est l'algorithme de Metropolis-Hastings. Il s'agit d'un algorithme simple, général et qui impose peu de restrictions. Il est détaillé dans les sections subséquentes du présent document.

Le rapport a la structure suivante. Dans la section 2, les informations pertinentes à l'étude sont présentées. On y retrouve une brève discussion sur les données et une introduction aux éléments de l'estimation bayésienne. La section 3 fait état des études déjà réalisées. Bien que les études bayésiennes soient encore dans une quantité négligeable, les études portant sur la modélisation des changements de prix d'actifs apportent des contributions intéressantes. La section 4 suit avec une analyse théorique du type de modèle choisi pour décrire les changements de prix et des hypothèses qui en découlent. L'analyse empirique de ce modèle constitue la section 5. Les statistiques descriptives des données sont présentées et les résultats obtenus par l'estimation bayésienne sont alors comparés à ceux obtenus par la méthode d'estimation classique du maximum de vraisemblance (MLE). Finalement, la section 6 apporte une conclusion sur les sujets exposés dans le rapport.

## 2 Informations pertinentes

### 2.1 Données

Au niveau des données, on constate que le nombre de transactions de titres est beaucoup plus important au cours des dernières années. Par exemple, le nombre de transactions était six fois plus élevé en 1999 qu'en 1991. On remarque aussi que la qualité des données à haute fréquence est moindre que celles à basse fréquence en raison du nombre non négligeable d'observations manquantes et de la précision douteuse dans l'enregistrement des transactions. Dans un autre ordre d'idée, la majorité des transactions dans une même journée n'impliquent pas un changement de prix du titre. Si on définit  $y_i \equiv P_{it} - P_{it-1}$ , alors  $y_i$  prend principalement la valeur zéro et d'autres valeurs discrètes. Campbell, Lo et MacKinlay (1997, p.122) suggèrent ainsi l'adoption d'un modèle non linéaire pour expliquer les changements de prix. L'estimation bayésienne d'un modèle non linéaire semble une méthode appropriée car le recours aux lois asymptotiques n'est pas nécessaire.

### 2.2 Éléments de l'analyse bayésienne

La suite de cette section est dédiée à une introduction aux éléments de l'approche bayésienne. Contrairement à l'approche classique du MLE ou des moindres carrés ordinaires (OLS) où les paramètres sont fixes et déterministes, la méthode bayésienne considère les paramètres comme étant stochastiques. Dans ce contexte, on utilise plutôt une distribution probabiliste des paramètres pour tenir compte de l'incertitude. L'approche consiste à combiner une distribution à priori selon un ensemble de croyances représentatives avec les données pour obtenir une distribution à posteriori sur laquelle on pratique l'inférence statistique.

En définissant  $\theta$  comme un vecteur  $k \times 1$  de paramètres inconnus et  $y$  étant les données, l'approche se résume par le théorème de Bayes suivant :

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(y|\theta)p(\theta)}{p(y)} \quad (2.1)$$

où  $p(\theta|y)$  est la distribution à posteriori,  $p(y|\theta)$  est la fonction de vraisemblance,  $p(\theta)$  est la distribution à priori et  $p(y)$  est la vraisemblance marginale.

Puisque cette dernière ne dépend pas de  $\theta$ , elle est constante et le théorème se réduit à :

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad (2.2)$$

Habituellement, le chercheur choisit une distribution à priori assez vague ne connaissant pas les valeurs possibles de  $\theta$  et laisse les données ajuster la distribution à posteriori. Avec beaucoup de données, on obtient une distribution à posteriori beaucoup plus précise que celle à priori.

Plusieurs objectifs de l'analyse bayésienne se résument à calculer :

$$E[g(\theta)|y] = \int_{\theta} g(\theta)p(\theta|y)d\theta \quad (2.3)$$

pour n'importe quelle fonction  $g(\theta)$  où l'espérance existe.

### 2.3 Algorithme de Metropolis-Hastings

Il n'est pratiquement jamais possible de calculer (2.3) de façon analytique. On a donc recours à des simulations. Il existe quelques algorithmes de simulation qui relèvent de la méthode MCMC. Les méthodes MCMC sont un cas particulier des méthodes Monte Carlo et sont employées pour simuler la distribution à posteriori quand la fonction de vraisemblance et la distribution à priori n'ont pas la même forme.

Les méthodes Monte Carlo ont en commun de tirer  $\theta^{(m)}$  pour  $m=1,\dots,M$ , un échantillon aléatoire provenant de  $p(\theta|y)$ , et d'utiliser le fait que:

$$\frac{1}{M} \left( \sum_{m=1}^M g(\theta^{(m)}) \right) \xrightarrow{p.s.} E[g(\theta)|y] \quad (2.4)$$

Dans le cas particulier MCMC, on recherche un processus markovien  $p(\theta^{(m+1)}|\theta^{(m)})$  pour lequel  $p(\theta|y)$  est la loi stationnaire unique. Il s'en suit que :

$$\theta^{(m)} \sim p(\theta|y) \Rightarrow \theta^{(m+1)} \sim p(\theta|y) \quad (2.5)$$

On applique ensuite (2.4) sous quelques conditions techniques. L'échantillonnage de Gibbs et la chaîne indépendante de Metropolis-Hastings sont deux algorithmes de simulation MCMC fréquemment utilisés.

Cependant, ces deux algorithmes imposent des restrictions très exigeantes. Ainsi, on a souvent recours à la marche aléatoire Metropolis-Hastings, un cas particulier de M-H, qui s'applique à tous les modèles non linéaires et à n'importe quelle distribution à priori. Il s'agit donc d'un outil de simulation flexible et c'est pourquoi il est utilisé dans ce rapport. Il consiste à trouver une chaîne de Markov dont la loi invariante est la loi à posteriori. Koop (2003, p.93) résume la démarche à suivre en quelques étapes :

- Choisir une valeur initiale  $\theta^{(0)}$ ;
- Générer un candidat selon :

$$\theta^* = \theta^{(m-1)} + z \quad (2.6)$$

où  $z$  est une variable aléatoire incrémentale avec une densité symétrique. La densité choisie est habituellement la normale multivariée  $f_N(\theta|\theta^{(m-1)}, \Sigma)$  avec moyenne  $\theta^{(m-1)}$  et matrice de covariance déterminée par le chercheur;

- S'assurer que la chaîne se déplace dans la bonne direction en suivant la règle d'acceptation du candidat avec la probabilité suivante :

$$\alpha(\theta^{(m-1)}, \theta^*) = \min \left[ \frac{p(\theta = \theta^* | y)}{p(\theta = \theta^{(m-1)} | y)}, 1 \right] \quad (2.7)$$

Pour appliquer la règle, il faut générer une valeur aléatoire tirée de loi uniforme (0,1) et la comparer à la valeur obtenue en (2.7). Si la valeur en (2.7) est supérieure à la valeur aléatoire, alors le candidat est accepté et  $\theta^{(m)} = \theta^*$ , sinon le candidat est rejeté et  $\theta^{(m)} = \theta^{(m-1)}$ ;

- Répéter les trois étapes précédentes M fois;
- Appliquer (2.4) afin d'approximer  $E[g(\theta)|y]$  pour n'importe quelle fonction  $g(\theta)$  où l'espérance existe.

Le point focal de l'algorithme est la règle d'acceptation du candidat. Une chaîne qui n'explore que les régions de haute probabilité à posteriori n'est pas désirable. Elle doit aussi visiter les régions de faible probabilité, mais à une fréquence moindre que celles de haute probabilité. Ainsi, la chaîne va souvent se déplacer des régions de faible à haute probabilité.

Dans la prochaine section, une revue des études antérieures sur la modélisation des changements de prix est présentée et l'analyse bayésienne est appliquée à l'un de ces modèles dans les sections 4 et 5.

### 3 Revue de la littérature

Malgré le peu d'études bayésiennes publiées, un certain nombre de modèles pour les changements de prix d'actifs échangés à haute fréquence sont suggérés dans la littérature. Les deux modèles les plus importants et discutés dans cette section sont le probit ordonné de Hausman, Lo et MacKinlay (1992) et le modèle de décomposition en trois composantes ADS de Rydberg et Shephard (1998). Dans un contexte plus général, Engle et Russell (1998) ont appliqué un modèle autorégressif conditionnel aux durées entre les transactions. McCulloch et Tsay (2001) ont développé un modèle bivarié pour les changements de prix et la durée entre les transactions.

#### 3.1 Le modèle probit ordonné

L'application du modèle probit ordonné par Hausman, Lo et Mackinlay (1992) se limite à l'étude d'un seul actif à la fois. Ils considèrent  $y_i^* = P_{t_i}^* - P_{t_{i-1}}^*$  le changement de prix de l'actif  $i$  où  $P_t^*$  est le prix virtuel de l'actif au temps  $t$ . Le modèle assume  $y_i^*$  comme une variable aléatoire continue et inobservable<sup>1</sup> suivant le modèle :

$$y_i^* = x_i \beta + \varepsilon_i \quad (3.1)$$

où  $x_i$  est un vecteur ligne de dimension- $p$  de variables explicatives au temps  $t_{i-1}$ ,  $\beta$  est un vecteur  $k \times 1$  de paramètres. Ils posent les hypothèses que  $E[\varepsilon_i | x_i] = 0$ ,  $Var[\varepsilon_i | x_i] = \sigma_i^2$  et  $Cov[\varepsilon_i, \varepsilon_j] = 0, \forall i \neq j$ . La variance conditionnelle  $\sigma_i^2$  est une fonction positive de la variable explicative  $w_i$  :

$$\sigma_i^2 = g(w_i) \quad (3.2)$$

Finalement, le changement de prix discret et observable  $y_i$  peut prendre  $k$  différentes valeurs. Les chercheurs doivent grouper les données selon l'ampleur du changement. Ils

---

<sup>1</sup> Le changement de prix observé est discret

mesurent l'ampleur du changement de prix en huitièmes de dollar et fixent  $k = 9$  ( $-4/8$  \$,  $-3/8$  \$, ...,  $+3/8$  \$,  $+4/8$  \$). On a donc la relation suivante entre  $y_i$  et  $y_i^*$  :

$$y_i = s_j \text{ si } \alpha_{j-1} < y_i^* \leq \alpha_j, j = 1, \dots, k \quad (3.3)$$

où les  $\alpha_j$  sont des nombres réels qui satisfont  $-\infty = \alpha_0 < \alpha_1 < \dots < \alpha_k = \infty$ . Les variables explicatives choisies sont la durée entre la transaction présente et la précédente, les trois premiers retards dans les changements de prix, l'écart entre le prix d'achat et le prix demandé par les «marketmakers» à la transaction précédente, les trois derniers retards dans la quantité d'actions échangées, le rendement de l'index Standard and Poor's 500 et une variable indicatrice. Les paramètres sont estimés par le maximum de vraisemblance.

Ils tirent trois principales conclusions : la durée entre les transactions affecte la moyenne et la variance conditionnelle des changements de prix, les coefficients des variables de retards des changements de prix sont négatifs et très significatifs et l'écart entre le prix d'achat et le prix demandé par les «marketmakers» a un effet sur la variance conditionnelle.

Bien que le modèle probit ordonné fonctionne bien pour ce type de données, le modèle développé par les auteurs est très lourd en raison du grand nombre de paramètres à estimer et ne permet pas de comprendre les fondements de la dépendance dynamique entre les changements de prix. De plus, le modèle nécessite un groupement des données selon l'ampleur des changements de prix.

### 3.2 Le modèle de décomposition en trois composantes ADS

Une alternative à l'approche de ces auteurs est le modèle de décomposition en trois composantes ADS de Rydberg et Shephard (1998). Il occupe une place importante dans cette section puisque l'analyse bayésienne est appliquée à ce modèle dans le cadre de cette recherche.

Le modèle ADS se limite lui aussi à l'étude d'un actif à la fois. Les trois composantes sont l'indicateur pour un changement de prix (A), la direction de ce changement (D) et son ampleur (S). Ils définissent le changement de prix à la  $i^{\text{ème}}$  transaction  $y_i$  :

$$y_i \equiv P_{ii} - P_{i-1} = A_i D_i S_i, \quad (3.4)$$

La variable  $A_i$  peut prendre seulement deux valeurs : 0, 1, la valeur 1 indiquant un changement de prix. Si  $A_i = 0$ , alors  $D_i = S_i = 0$ . Si  $A_i = 1$ , alors la variable  $D_i$  prend la valeur 1 si le prix augmente à la transaction  $i$  et prend la valeur -1 si le prix diminue. La variable  $S_i$  prendre les valeurs suivantes : 1, 2, 3, ..., des entiers positifs et multiples du plus petit changement (1/8 de dollar) pour quantifier l'ampleur du changement de prix à la transaction  $i$ . En définissant  $F_i$  l'information disponible à la transaction  $i$ , ils partitionnent l'évolution des changements de prix sous (3.4) en trois blocs :

$$P(y_i | F_{i-1}) = P(S_i | D_i, A_i, F_{i-1}) P(D_i | A_i, F_{i-1}) P(A_i | F_{i-1}) \quad (3.5)$$

Ce choix permet aux auteurs de modéliser chacune des trois composantes séparément de la façon suivante :

- $P(A_i | F_{i-1})$  comme un processus binaire sur  $\{0,1\}$  modélisant l'activité dans le changement de prix, s'il change ou pas;
- $P(D_i | A_i, F_{i-1})$  comme un autre processus binaire sur  $\{-1,1\}$  modélisant la direction du changement de prix;
- $P(S_i | D_i, A_i, F_{i-1})$  comme un processus sur des entiers positifs quantifiant l'ampleur du changement de prix.

Pour les deux premières composantes, ils choisissent un modèle non linéaire logit pour les processus binaires. Ils considèrent l'évolution de la probabilité  $p_i = P(A_i = 1)$  dans le temps en assumant que :

$$\ln\left(\frac{p_i}{1-p_i}\right) = x_i \beta \quad (3.6)$$

où  $x_i$  est un vecteur d'une dimension finie constitué d'éléments de  $F_{i-1}$  et  $\beta$  un vecteur de paramètres. Si  $A_i = 1$ , alors  $D_i$  est une variable binaire et les auteurs considèrent l'évolution de la probabilité  $\delta_i = P(D_i = 1 | A_i = 1)$  dans le temps en assumant que :

$$\ln\left(\frac{\delta_i}{1-\delta_i}\right) = z_i \gamma \quad (3.7)$$

où  $z_i$  est un vecteur d'une dimension finie constitué d'éléments de  $F_{i-1}$  et  $\gamma$  un vecteur de paramètres. Pour modéliser la troisième composante  $S_i$ , les auteurs tiennent compte de l'asymétrie entre les changements de prix positifs et négatifs. Pour ce faire, ils assument que :

$$S_i | (D_i = 1, A_i = 1) \sim 1 + g(\lambda_{u,i}) \quad (3.8)$$

$$S_i | (D_i = -1, A_i = 1) \sim 1 + g(\lambda_{d,i})$$

où  $g(\lambda)$  est une distribution géométrique avec paramètre  $\lambda$ . Les paramètres  $\lambda_{j,i}$  évoluent dans le temps selon :

$$\ln\left(\frac{\lambda_{j,i}}{1-\lambda_{j,i}}\right) = w_i \theta_j, j = u, d \quad (3.9)$$

où  $w_i$  est un vecteur d'une dimension finie constitué d'éléments de  $F_{i-1}$  et  $\theta_j$  un vecteur de paramètres. Ces modèles permettent de classer la transaction  $i$  dans l'une des trois catégories suivantes :

- Aucun changement de prix :  $A_i = 0$  avec probabilité  $(1 - p_i)$ ;
- Une augmentation du changement de prix :  $A_i = 1, D_i = 1$  avec probabilité  $p_i \delta_i$  et une ampleur du changement suivant  $1 + g(\lambda_{u,i})$ ;
- Une diminution du changement de prix :  $A_i = 1, D_i = -1$  avec probabilité  $p_i (1 - \delta_i)$  et une ampleur du changement suivant  $1 + g(\lambda_{d,i})$ .

Les auteurs choisissent comme éléments de  $F_{i-1}$  (variables explicatives) l'activité du changement de prix à la dernière transaction pour (3.6), la direction du changement à la dernière transaction pour (3.7) et l'ampleur du dernier changement pour (3.9). Ils ont tenté aussi d'inclure la quantité d'actions échangées à la dernière transaction, la durée entre les deux dernières transactions et l'écart entre le prix d'achat et le prix demandé par les «marketmakers». Elles se sont toutefois avérées non significatives. Les paramètres  $\beta, \gamma, \theta_u, \theta_d$  sont estimés séparément par la méthode du MLE.

Les auteurs ont réussi avec ce modèle de décomposition ADS à démontrer la dépendance dynamique dans les changements de prix. Ils constatent que les changements de prix se produisent en groupe car la probabilité qu'il y ait un changement à la transaction  $i$  sachant qu'il y a eu changement à  $i-1$  est deux fois plus élevée. Ils remarquent aussi que la direction du changement a tendance à s'inverser d'une transaction à l'autre et que la probabilité d'un changement d'une grande ampleur est proportionnelle à l'ampleur du changement précédent.

La principale faiblesse de l'étude est le peu de variables explicatives significatives. Aucun facteur externe comme la durée ou l'index S&P 500 n'est utilisé. Dans un contexte bayésien, de telles variables peuvent s'avérer importantes. De plus, tout comme le modèle de Hausman, Lo et MacKinlay (1992), le modèle est limité à l'étude d'un seul actif à la fois.

### 3.3 Études bayésiennes

Dans les études bayésiennes antérieures reliées à la finance, on retrouve entre autres Jacquier, Polson et Rossi (1994) qui traitent de modèles à volatilité stochastique, Li (1999) de la durée entre transactions et Darsinos et Satchell (2001) des prix d'options. L'inconvénient majeur des études bayésiennes est le calcul computationnel parfois très lourd. Le choix de la loi à priori entraîne aussi son lot de difficultés dans certains cas.

Cette section a fait état des principales études antérieures. Le modèle probit avait comme lacune première de ne pas pouvoir démontrer la dépendance dynamique dans les changements de prix, ce qui a été corrigé dans le modèle de décomposition ADS. Cependant, ce modèle ne permet pas d'inclure différentes variables explicatives dans un contexte d'estimation classique du MLE. Dans la prochaine section, le modèle ADS est repris et intégré dans une approche bayésienne. Des variables explicatives sont ajoutées pour tenter d'enrichir le modèle.

## 4 Analyse théorique

Dans cette section, le but est d'intégrer le modèle ADS discuté dans la section 3 dans le contexte bayésien. Ainsi, chacun des modèles (3.6), (3.7) et (3.9) est estimé séparément par l'analyse bayésienne. On doit retrouver une forme similaire à (2.2) afin de simuler la distribution à posteriori avec l'algorithme de Metropolis-Hastings. Pour y arriver, on doit déterminer la fonction de vraisemblance et fixer une loi à priori.

### 4.1 La vraisemblance du modèle ADS

Le modèle (3.6) est un logit et sa fonction de vraisemblance est déterminée de la façon suivante<sup>2</sup>:

$$\begin{aligned}
 L(p_i; a_1, a_2, \dots, a_n) &= P(A_1 = a_1)P(A_2 = a_2) \dots P(A_n = a_n) \\
 &= \prod_{i=1}^n P(A_i = a_i) \\
 &= \prod_{i=1}^n p_i^{a_i} (1 - p_i)^{1-a_i}
 \end{aligned}$$

---

<sup>2</sup> Pour plus de détails, les lecteurs intéressés peuvent consulter Wooldridge (2003, p.557)

$$= \prod_{i=1}^n \left( \frac{p_i}{1-p_i} \right)^{a_i} (1-p_i) \quad (4.1)$$

En réorganisant les termes dans (3.6), on obtient :

$$1-p_i = \frac{p_i}{e^{x_i\beta}} = \frac{1}{1+e^{x_i\beta}}$$

En insérant ce résultat dans (4.1), on obtient la fonction de vraisemblance :

$$L(A_1, A_2, \dots, A_n | \beta) = \prod_{i=1}^n \left( e^{x_i\beta} \right)^{A_i} \left( \frac{1}{1+e^{x_i\beta}} \right)$$

Il est plus convenable de l'exprimer sous forme logarithmique :

$$\log L(A_1, A_2, \dots, A_n | \beta) = \sum_{i=1}^n [A_i x_i \beta - \log(1+e^{x_i\beta})] \quad (4.2)$$

En suivant les mêmes démarches que pour (3.6), on trouve la fonction de vraisemblance pour (3.7) sous forme logarithmique :

$$\log L(D_1, D_2, \dots, D_n | \gamma) = \sum_{i=1}^n [D_i z_i \gamma - \log(1+e^{z_i\gamma})] \quad (4.3)$$

On sait par (3.8) que  $\lambda_i(1-\lambda_i)^{s_i-1} = P(S_i = s_i | D_i = d_i, A_i = 1)$ .

La fonction de vraisemblance pour (3.9) est déterminée ainsi :

$$\begin{aligned} L(\lambda_i; s_1, s_2, \dots, s_n) &= P(S_1 = s_1)P(S_2 = s_2)\dots P(S_n = s_n) \\ &= \prod_{i=1}^n P(S_i = s_i) \\ &= \prod_{i=1}^n \lambda_i(1-\lambda_i)^{s_i-1} \\ &= \prod_{i=1}^n \left( \frac{\lambda_i}{1-\lambda_i} \right) (1-\lambda_i)^{s_i} \end{aligned} \quad (4.4)$$

En réorganisant les termes dans (3.9), on obtient :

$$1-\lambda_i = \frac{\lambda_i}{e^{w_i\theta_j}} = \frac{1}{1+e^{w_i\theta_j}}$$

En insérant ce résultat dans (4.4), on obtient la fonction de vraisemblance :

$$L(S_1, S_2, \dots, S_n | \theta_j) = \prod_{i=1}^n \left( e^{w_i\theta_j} \right) \left( \frac{1}{1+e^{w_i\theta_j}} \right)^{S_i}$$

Sous forme logarithmique, on obtient :

$$\log L(S_1, S_2, \dots, S_n | \theta_j) = \sum_{i=1}^n [w_i \theta_j - S_i \log(1 + e^{w_i \theta_j})] \quad (4.5)$$

Les fonctions de vraisemblance calculées pour chacun des trois modèles, il reste à choisir une distribution à priori pour les paramètres  $\beta, \gamma, \theta_u, \theta_d$ .

## 4.2 Le choix de la loi à priori

Si l'on veut obtenir une distribution à posteriori plus raffinée que celle à priori, il est recommandé de choisir une loi à priori qui n'a pas beaucoup d'influence sur la loi à posteriori. Ainsi, le choix de la loi à priori pour tous les paramètres s'arrête sur la loi normale multivariée  $N(0, \Sigma)$  avec une variance élevée afin d'affaiblir le plus possible l'influence de la loi. De plus, on fixe les termes de covariance à zéro. La loi à priori choisie est la même pour tous les modèles car les résultats obtenus sont robustes aux changements de cette loi :

$$\Sigma_\beta = \Sigma_\gamma = \Sigma_{\theta_j} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}, \quad j = u, d$$

Avant de passer à la construction de l'algorithme de la marche aléatoire M-H, on peut exprimer la densité à posteriori pour chacun des trois modèles en enlevant les termes qui n'impliquent pas les paramètres :

$$\log p(\beta | A_i) \propto \sum_{i=1}^n [A_i x_i \beta - \log(1 + e^{x_i \beta})] - \frac{1}{2} (\beta' \Sigma_\beta^{-1} \beta) \quad (4.6)$$

$$\log p(\gamma | D_i) \propto \sum_{i=1}^n [D_i z_i \gamma - \log(1 + e^{z_i \gamma})] - \frac{1}{2} (\gamma' \Sigma_\gamma^{-1} \gamma) \quad (4.7)$$

$$\log p(\theta_j | S_i) \propto \sum_{i=1}^n [w_i \theta_j - S_i \log(1 + e^{w_i \theta_j})] - \frac{1}{2} (\theta_j' \Sigma_{\theta_j}^{-1} \theta_j), \quad j = u, d \quad (4.8)$$

À partir des distributions à posteriori, il est possible de simuler les valeurs des paramètres avec la marche aléatoire Metropolis-Hastings. Dans la prochaine section, cette approche bayésienne est mise de l'avant pour tester les modèles. Plusieurs variables explicatives sont utilisées pour tenter d'enrichir le modèle ADS développé par Rydberg et Shephard (1998).

## 5 Analyse empirique

Cette section traite des données utilisées, des résultats des estimations obtenues par la méthode classique du maximum de vraisemblance et par l'approche bayésienne. Elle comprend également une discussion sur les forces et faiblesses de chacun des modèles estimés par l'analyse bayésienne.

### 5.1 Ajustement des données

Au départ, on dénombrait 59 838 observations sur une période de 63 jours de transactions. En se servant de variables de premier retard, on conditionne toujours sur la première transaction de chaque journée. Ainsi, la taille réelle de l'échantillon est réduite à 59 775 observations. Dans le cadre de l'analyse empirique bayésienne, le nombre de données utilisées a été considérablement diminué pour des raisons computationnelles. Ainsi, l'estimation des paramètres est réalisée avec un sous échantillon comprenant les 25 000 premières observations des 59 775. Néanmoins, les autres observations regroupées dans un deuxième sous échantillon demeurent utiles pour mesurer la qualité des prévisions des valeurs de ce sous échantillon sachant le premier. Avant de présenter les résultats des estimations obtenus par l'approche bayésienne, il convient de discuter de ceux obtenus par l'approche classique du maximum de vraisemblance.

## 5.2 Résultats des estimations par MLE

L'estimation des paramètres du modèle ADS par maximum de vraisemblance a été faite au préalable par Rydberg et Shephard (1998) dans leur étude. Les variables explicatives considérées étaient :

1.  $A_{i-1}$  : Indicatrice d'un changement de prix à la transaction précédente;
2.  $D_{i-1}$  : Indicatrice de la direction du changement de prix à la transaction précédente;
3.  $S_{i-1}$  : Ampleur du changement de prix à la transaction précédente;
4.  $V_{i-1}$  : Le volume échangé de l'action à la dernière transaction, divisé par 1000;
5.  $\Delta t_{i-1}$  : La durée entre les transactions (i-2) et (i-1);
6.  $BA_i$  : L'écart entre les prix de vente et d'achat de l'action induit par les «market makers». Cette information est connue au moment de la transaction i.

Les variables explicatives 4, 5 et 6 se sont avérées non significatives. Les auteurs ont donc seulement utilisé les variables de premier retard, ce qui donne le modèle de décomposition ADS suivant :

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 A_{i-1} \quad (5.1)$$

$$\ln\left(\frac{\delta_i}{1-\delta_i}\right) = \gamma_0 + \gamma_1 D_{i-1} \quad (5.2)$$

$$\ln\left(\frac{\lambda_{u,i}}{1-\lambda_{u,i}}\right) = \theta_{u,0} + \theta_{u,1} S_{i-1} \quad (5.3)$$

$$\ln\left(\frac{\lambda_{d,i}}{1-\lambda_{d,i}}\right) = \theta_{d,0} + \theta_{d,1} S_{i-1} \quad (5.4)$$

Les paramètres  $\beta, \gamma, \theta_u, \theta_d$  sont estimés séparément par le maximum de vraisemblance et les résultats obtenus pour l'échantillon de 59 775 observations sont les suivants :

*Tableau 1: Estimation des paramètres du modèle ADS par maximum de vraisemblance pour l'action IBM échangée entre le 01/11/90 et le 31/01/91*

<b>Paramètre</b>	$\beta_0$	$\beta_1$	$\gamma_0$	$\gamma_1$	$\theta_{u,0}$	$\theta_{u,1}$	$\theta_{d,0}$	$\theta_{d,1}$
<b>Estimé</b>	-1,057	0,962	-0,067	-2,307	2,235	-0,670	2,085	-0,509
<b>Écart-type</b>	0,104	0,044	0,023	0,056	0,029	0,050	0,187	0,139

*Source : Rydberg et Shephard (1998)*

Les variables de premier retard étant toutes statistiquement significatives, l'interprétation de ces résultats par Tsay (2002, p.193) est particulièrement intéressante. L'estimation des paramètres du modèle (5.1) indique une tendance à observer des changements de prix consécutifs :

$$P(A_i = 1 | A_{i-1} = 0) = 0,258 \quad P(A_i = 1 | A_{i-1} = 1) = 0,476$$

En effet, la probabilité d'observer un changement de prix à la transaction  $i$  est presque deux fois plus élevée quand il y a eu un changement de prix à la transaction  $i-1$ .

L'estimation des paramètres du modèle (5.2) suggère que la probabilité d'observer des augmentations ou diminutions de prix consécutives est très faible. De plus, les probabilités d'une augmentation ou d'une diminution de prix à la transaction  $i$  sachant qu'aucun changement de prix n'est survenu à la transaction  $i-1$  sont pratiquement les mêmes :

$$P(D_i = 1 | A_{i-1} = 0) = 0,483 \quad P(D_i = 1 | A_{i-1} = 1, D_{i-1} = 1) = 0,085$$

$$P(D_i = 1 | A_{i-1}, D_{i-1} = -1) = 0,904$$

Ces probabilités révèlent des inversions fréquentes dans la direction des changements de prix d'actions échangées à haute fréquence.

Finalement, l'estimation des paramètres des modèles (5.3) et (5.4) tend à démontrer qu'un grand changement de prix a une probabilité plus élevée d'être suivi par un autre grand changement. Pour les changements positifs, on remarque que la probabilité d'observer une augmentation de prix de 1/8 de dollar ( $S_i = 1$ ) diminue au fur et à mesure que le changement précédent ( $S_{i-1}$ ) est grand :

$$P(S_i = 1 | S_{i-1} = 1, D_i = 1) = 0,827$$

$$P(S_i = 1 | S_{i-1} = 2, D_i = 1) = 0,709$$

$$P(S_i = 1 | S_{i-1} = 3, D_i = 1) = 0,556$$

Ainsi, la probabilité d'observer un grand  $S_i$  est proportionnelle à  $S_{i-1}$ , sachant qu'il y a une augmentation du prix de l'action à la transaction  $i$ .

### 5.3 Résultats des estimations par l'analyse bayésienne

Dans le contexte d'estimation par MLE, les modèles (5.1) à (5.4) ne contiennent qu'une seule variable explicative significative, soit la variable de premier retard. Avec l'approche bayésienne, les premiers modèles estimés comprennent seulement une constante. Par la suite, la variable de premier retard et les variables  $BA_i$  et  $V_{i-1}$  sont successivement ajoutées. La variable  $\Delta t_{i-1}$  a été abandonnée car elle n'apporte aucune information pertinente pour chacun des modèles étudiés. Le nombre d'itérations dans chacune des simulations est de  $M=30\ 000$ . Les résultats présentés pour chacune des composantes A, D et S sont les suivants :

- La moyenne ( $\bar{\theta}$ ), l'écart-type ( $\bar{\sigma}$ ), l'asymétrie ( $\bar{S}$ ) et l'aplatissement ( $\bar{K}$ ) de la distribution à posteriori des paramètres;
- La moyenne des probabilités d'acceptation des candidats ( $\overline{\alpha(\theta^{(m-1)}, \theta^*)}$ ) dans la simulation marche aléatoire de Metropolis-Hastings;
- La moyenne des vraisemblances prédictives ( $\overline{L(y_i^{(B)} | y_i^{(A)}, \theta^{(A)})}$ ), une mesure de la qualité des prévisions.

Les quatre premiers moments de la distribution à posteriori sont calculés ainsi :

$$\bar{\theta} = \frac{1}{M} \sum_{m=1}^M \theta^{(m)} \quad (5.5)$$

$$\hat{\sigma} = \sqrt{\frac{1}{M} \sum_{m=1}^M (\theta^{(m)} - \bar{\theta})^2} \quad (5.6)$$

$$\hat{S} = \frac{1}{M\hat{\sigma}^3} \sum_{m=1}^M (\theta^{(m)} - \bar{\theta})^3 \quad (5.7)$$

$$\hat{K} = \frac{1}{M\hat{\sigma}^4} \sum_{m=1}^M (\theta^{(m)} - \bar{\theta})^4 \quad (5.8)$$

La moyenne des probabilités d'acceptation est une mesure du déplacement de la chaîne au cours de la simulation. Une moyenne trop faible indique que la chaîne ne bouge pas souvent car les candidats sont plus souvent qu'autrement rejetés. À l'inverse, une moyenne trop élevée révèle qu'à chaque itération, les candidats ont des valeurs qui diffèrent peu. L'expérience dans la discipline révèle que la moyenne de la probabilité d'acceptation se situe autour de 0,5 et ne devrait pas être inférieure à 0,25 ou supérieure à 0,75. Elle est calculée de cette façon :

$$\overline{\alpha(\theta^{(m-1)}, \theta^*)} = \frac{1}{M} \sum_{m=1}^M \alpha(\theta^{(m-1)}, \theta^*)^{(m)} \quad (5.9)$$

La moyenne des vraisemblances prédictives est une mesure de la qualité des prévisions des valeurs du deuxième sous échantillon B (les observations  $y_i$  25 001 à 59 775) sachant le premier A (les paramètres estimés selon les premières 25 000 observations). Plus la valeur calculée est élevée, meilleure est la qualité des prévisions. Elle est obtenue comme en (5.9) :

$$\overline{L(y_i^{(B)} | y_i^{(A)}, \theta^{(A)})} = \frac{1}{M} \sum_{m=1}^M L(y_i^{(B)} | y_i^{(A)}, \theta^{(A)})^{(m)} \quad (5.10)$$

Afin de compléter les tableaux des résultats, quelques histogrammes des distributions à posteriori et traces des paramètres sont présentés.

### 5.3.1 Résultats pour la composante A

Pour la composante A, les modèles estimés successivement sont les suivants :

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 \quad (5.11)$$

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 A_{i-1} \quad (5.12)$$

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 A_{i-1} + \beta_2 BA_i \quad (5.13)$$

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 A_{i-1} + \beta_2 BA_i + \beta_3 V_{i-1} \quad (5.14)$$

Les modèles sont estimés séparément en considérant le sous échantillon des 25 000 premières transactions. Les résultats sont présentés dans les tableaux suivants :

*Tableau 2: Estimation bayésienne des paramètres du modèle (5.11) pour les 25 000 premières transactions de l'action IBM*

Variable	$\bar{\beta}$	$\hat{\sigma}_\beta$	$\hat{S}_\beta$	$\hat{K}_\beta$	$\bar{\alpha}(\beta^{(m-1)}, \beta^*)$	$\bar{L}(A_i^{(B)}   A_i^{(A)}, \beta^{(A)})$
Cte	-0,7502	0,0135	-0,0271	3,0059	0,4484	-22 216

Source : Estimations du chercheur<sup>3</sup>

<sup>3</sup> Les estimations sont faites avec le logiciel MATLAB 7.0

Tableau 3: Estimation bayésienne des paramètres du modèle (5.12) pour les 25 000 premières transactions de l'action IBM

Variable	$\bar{\beta}$	$\hat{\sigma}_{\beta}$	$\hat{S}_{\beta}$	$\hat{K}_{\beta}$	$\overline{\alpha(\beta^{(m-1)}, \beta^*)}$	$\overline{L(A_i^{(B)}   A_i^{(A)}, \beta^{(A)})}$
Cte	-1,0809	0,0176	0,0045	3,0358	0,3452	-21 365
$A_{i-1}$	0,9360	0,0286	-0,0253	3,0692		

Source : Estimations du chercheur

Tableau 4: Estimation bayésienne des paramètres du modèle (5.13) pour les 25 000 premières transactions de l'action IBM

Variable	$\bar{\beta}$	$\hat{\sigma}_{\beta}$	$\hat{S}_{\beta}$	$\hat{K}_{\beta}$	$\overline{\alpha(\beta^{(m-1)}, \beta^*)}$	$\overline{L(A_i^{(B)}   A_i^{(A)}, \beta^{(A)})}$
Cte	-1,0959	0,0264	-0,0296	2,9260	0,4254	-21 358
$A_{i-1}$	0,9357	0,0285	-0,0399	2,9602		
$BA_i$	0,0368	0,0488	-0,0271	2,8816		

Source : Estimations du chercheur

Tableau 5: Estimation bayésienne des paramètres du modèle (5.14) pour les 25 000 premières transactions de l'action IBM

Variable	$\bar{\beta}$	$\hat{\sigma}_{\beta}$	$\hat{S}_{\beta}$	$\hat{K}_{\beta}$	$\overline{\alpha(\beta^{(m-1)}, \beta^*)}$	$\overline{L(A_i^{(B)}   A_i^{(A)}, \beta^{(A)})}$
Cte	-1,0706	0,0282	-0,2387	3,3663	0,3165	-21 338
$A_{i-1}$	0,9379	0,0291	-0,0279	2,8412		
$BA_i$	0,0458	0,0503	-0,3928	3,9802		
$V_{i-1}$	-0,0172	0,0038	-0,1095	3,0647		

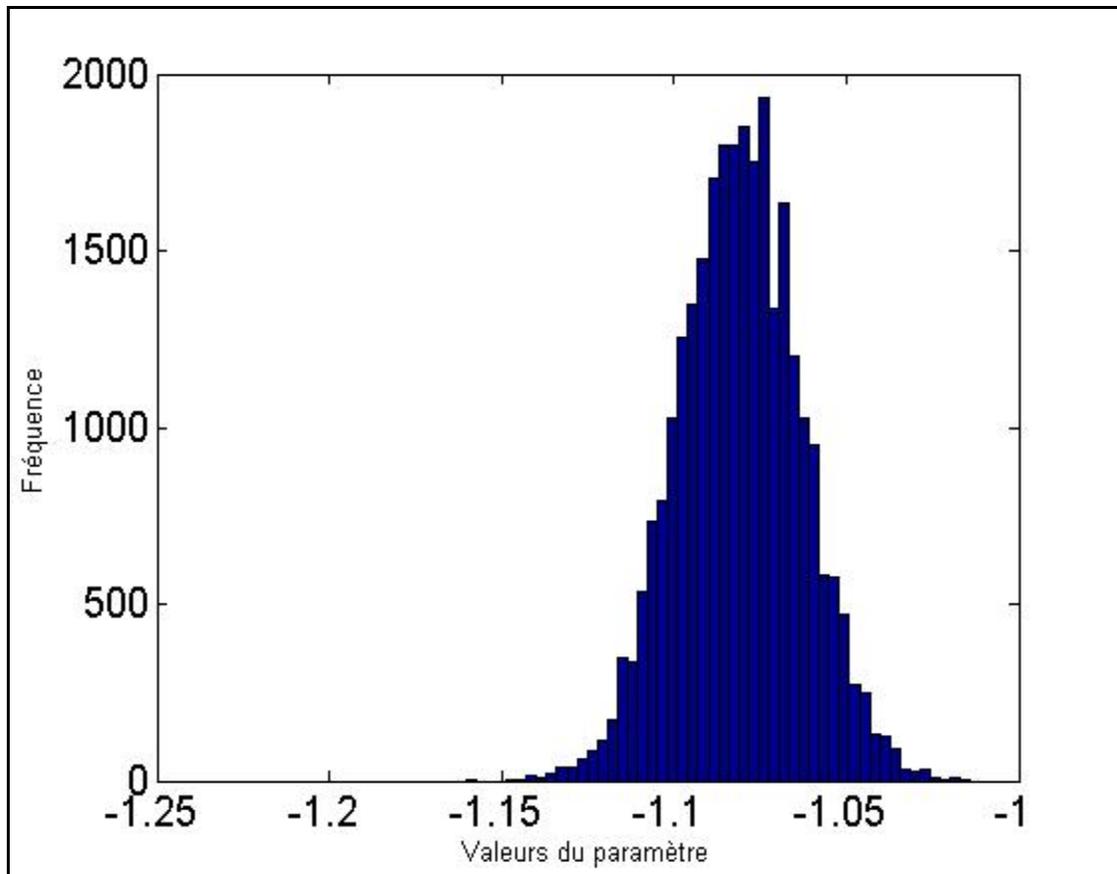
Source : Estimations du chercheur

Les valeurs obtenues pour le modèle (5.12) sont sensiblement les mêmes que celles estimées par le maximum de vraisemblance pour le modèle (5.1). Le fait d'inclure la variable  $BA_i$  n'a pas un impact statistique tangible car l'écart-type à posteriori de ce paramètre est élevé par rapport à sa moyenne à posteriori.

Cependant, la variable  $V_{i-1}$  apporte de l'information dans les modèles étudiés selon la moyenne des vraisemblances prédictives calculée. En effet, la variable  $V_{i-1}$  augmente substantiellement la qualité des prévisions pour la présence d'un changement de prix. Par ailleurs, les différents moments des distributions à posteriori des paramètres estimés semblent démontrer la normalité et la précision de ces distributions.

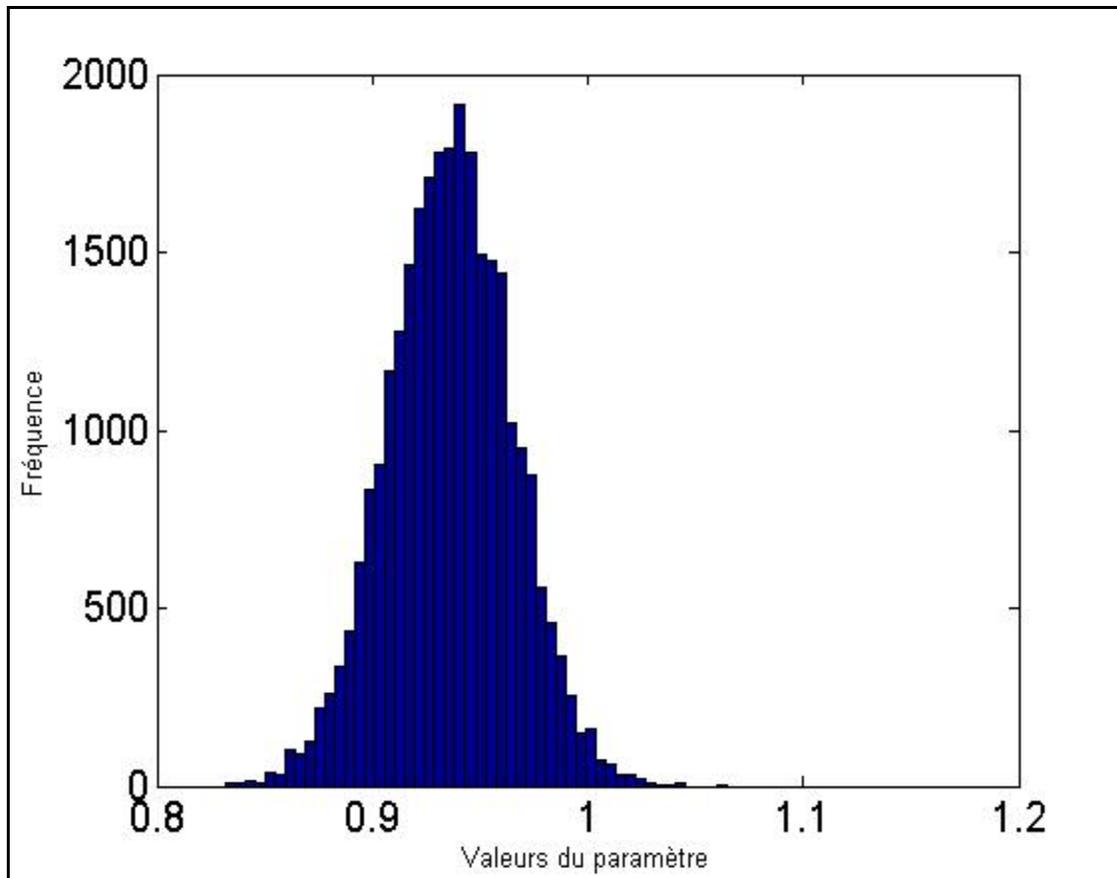
Les prochains graphiques reflètent les résultats présentés dans les tableaux précédents. Les histogrammes illustrent les distributions à posteriori des paramètres et les traces des paramètres montrent les variations des valeurs des paramètres tout au long de l'algorithme de la marche aléatoire M-H. Afin d'alléger le présent document, seulement les graphiques obtenus pour les paramètres du modèle (5.12) y sont inclus.

Graphique 1 : Distribution à posteriori du paramètre  $\beta_0$  dans le modèle (5.12) pour les 25 000 premières transactions de l'action IBM



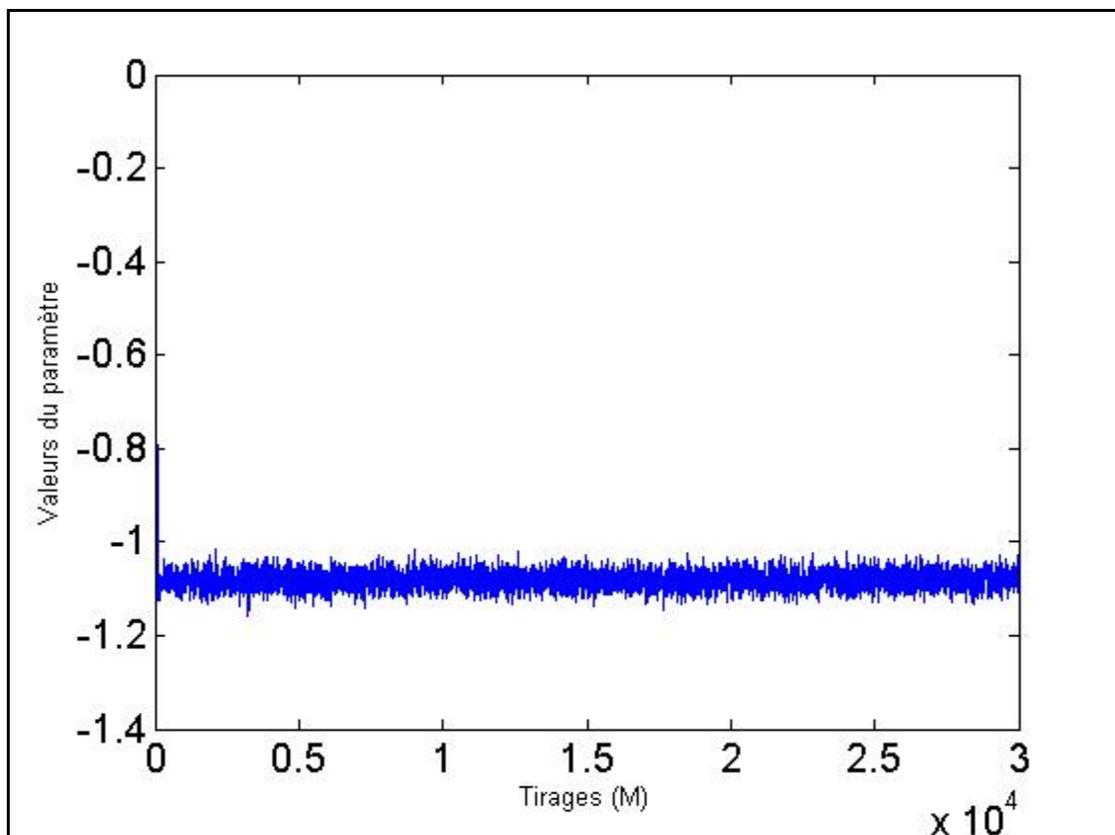
Source : Estimations du chercheur

Graphique 2 : Distribution à posteriori du paramètre  $\beta_1$  dans le modèle (5.12) pour les 25 000 premières transactions de l'action IBM



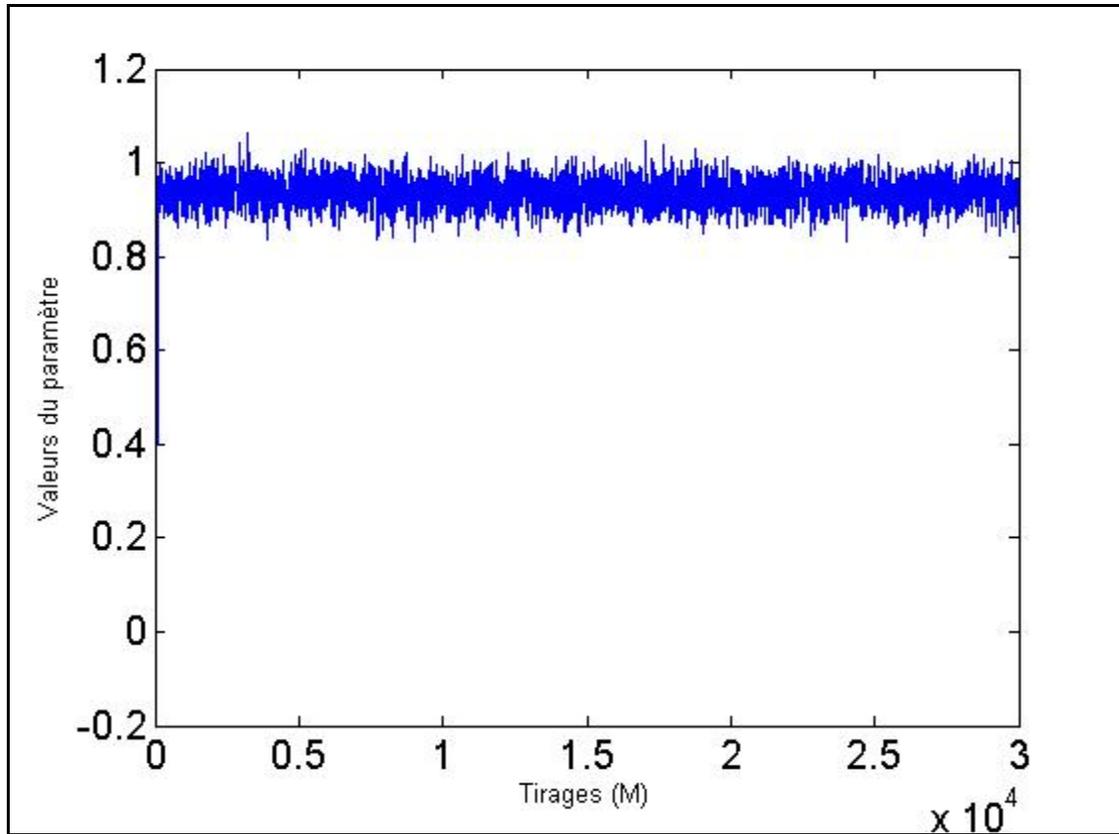
Source : Estimations du chercheur

Graphique 3 : Trace du paramètre  $\beta_0$  dans le modèle (5.12) lors de l'algorithme M-H pour les 25 000 premières transactions de l'action IBM



Source : Estimations du chercheur

Graphique 4 : Trace du paramètre  $\beta_1$  dans le modèle (5.12) lors de l'algorithme M-H pour les 25 000 premières transactions de l'action IBM



Source : *Estimations du chercheur*

Les traces observées indiquent clairement que la convergence vers la loi à posteriori est très rapide pour les deux paramètres. Afin d'éviter l'effet des valeurs obtenues avant la convergence, les 2 000 premiers tirages<sup>4</sup> ont été retirés des distributions pour toutes les estimations.

---

<sup>4</sup> Les premiers tirages retirés sont communément appelés «burn-in» dans la littérature

### 5.3.2 Résultats pour la composante D

Pour la composante D, les modèles estimés successivement sont les suivants :

$$\ln\left(\frac{\delta_i}{1-\delta_i}\right) = \gamma_0 \quad (5.15)$$

$$\ln\left(\frac{\delta_i}{1-\delta_i}\right) = \gamma_0 + \gamma_1 D_{i-1} \quad (5.16)$$

$$\ln\left(\frac{\delta_i}{1-\delta_i}\right) = \gamma_0 + \gamma_1 D_{i-1} + \gamma_2 BA_i \quad (5.17)$$

$$\ln\left(\frac{\delta_i}{1-\delta_i}\right) = \gamma_0 + \gamma_1 D_{i-1} + \gamma_2 BA_i + \gamma_3 V_{i-1} \quad (5.18)$$

Pour l'étude de cette composante D (direction), on n'utilise évidemment que les données pour lesquelles un changement de prix est observé. Des 25 000 premières transactions enregistrées, on en dénombre 8 019 avec un changement de prix. Les résultats sont présentés dans les tableaux suivants :

*Tableau 6 : Estimation bayésienne des paramètres du modèle (5.15) pour les 8 019 premières transactions avec changement de prix de l'action IBM*

Variable	$\bar{\gamma}$	$\hat{\sigma}_\gamma$	$\hat{S}_\gamma$	$\hat{K}_\gamma$	$\bar{\alpha}(\gamma^{(m-1)}, \gamma^*)$	$\bar{L}(D_i^{(B)}   D_i^{(A)}, \gamma^{(A)})$
Cte	0,0071	0,0226	0,0284	3,0080	0,6030	-8 109,6

*Source : Estimations du chercheur*

Tableau 7 : Estimation bayésienne des paramètres du modèle (5.16) pour les 8 019 premières transactions avec changement de prix de l'action IBM

Variable	$\bar{\gamma}$	$\hat{\sigma}_\gamma$	$\hat{S}_\gamma$	$\hat{K}_\gamma$	$\bar{\alpha}(\gamma^{(m-1)}, \gamma^*)$	$\bar{L}(D_i^{(B)}   D_i^{(A)}, \gamma^{(A)})$
Cte	-0,0624	0,0270	-0,0265	2,9158	0,5958	-5 902,1
$D_{i-1}$	-2,3037	0,0558	-0,0781	3,0410		

Source : Estimations du chercheur

Tableau 8 : Estimation bayésienne des paramètres du modèle (5.17) pour les 8 019 premières transactions avec changement de prix de l'action IBM

Variable	$\bar{\gamma}$	$\hat{\sigma}_\gamma$	$\hat{S}_\gamma$	$\hat{K}_\gamma$	$\bar{\alpha}(\gamma^{(m-1)}, \gamma^*)$	$\bar{L}(D_i^{(B)}   D_i^{(A)}, \gamma^{(A)})$
Cte	0,0088	0,0471	0,0694	2,9832	0,6670	-5 899,8
$D_{i-1}$	-2,3057	0,0576	-0,0069	2,9603		
$BA_i$	-0,1729	0,0966	-0,1196	3,0783		

Source : Estimations du chercheur

Tableau 9 : Estimation bayésienne des paramètres du modèle (5.18) pour les 8 019 premières transactions avec changement de prix de l'action IBM

Variable	$\bar{\gamma}$	$\hat{\sigma}_\gamma$	$\hat{S}_\gamma$	$\hat{K}_\gamma$	$\bar{\alpha}(\gamma^{(m-1)}, \gamma^*)$	$\bar{L}(D_i^{(B)}   D_i^{(A)}, \gamma^{(A)})$
Cte	0,0824	0,0468	-0,1406	2,9264	0,3068	-5 851,7
$D_{i-1}$	-2,2961	0,0549	0,0312	2,7875		
$BA_i$	-0,1767	0,0907	0,2197	2,9552		
$V_{i-1}$	-0,0462	0,0082	-0,1753	2,9629		

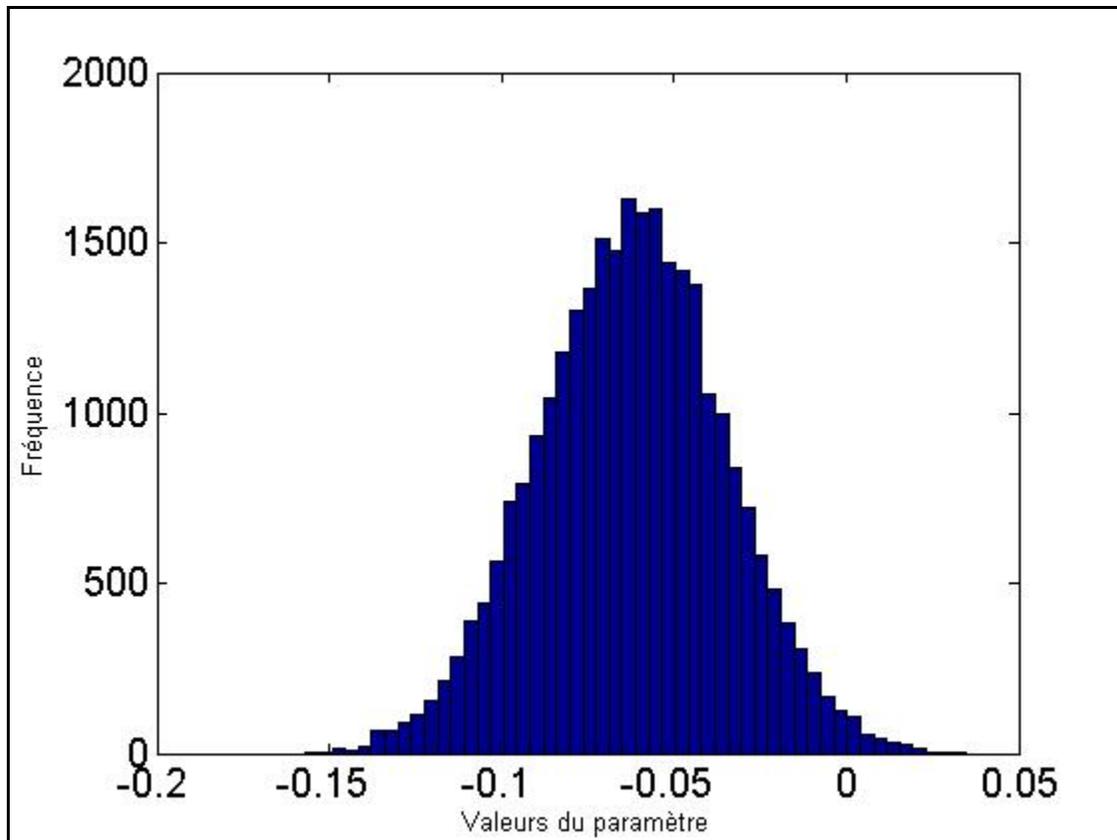
Source : Estimations du chercheur

Encore une fois, les valeurs estimées pour le modèle (5.16) sont pratiquement les mêmes que celles estimées par le maximum de vraisemblance pour le modèle (5.2). Le paramètre de la variable  $BA_i$  a une valeur moyenne négative mais a peu d'effet sur la probabilité d'observer une hausse ou une baisse du prix de l'actif. En conditionnant sur les valeurs moyennes des paramètres (voir le tableau 8), la probabilité d'une hausse de prix sachant qu'il n'y a pas eu de changement de prix à la dernière transaction est de 0,496 si  $BA_i = 1/8$  \$ (valeur minimale de l'échantillon) et de 0,481 si  $BA_i = 4/8$  \$ (valeur maximale la plus observée de l'échantillon). On calcule  $P(D_i = 1 | A_{i-1} = 0) = 0,483$  avec l'estimation du modèle (5.2) par maximum de vraisemblance.

Il est à noter que l'inclusion de la variable  $V_{i-1}$  dans le modèle (5.18) apporte une meilleure valeur prédictive pour la direction d'un changement de prix.

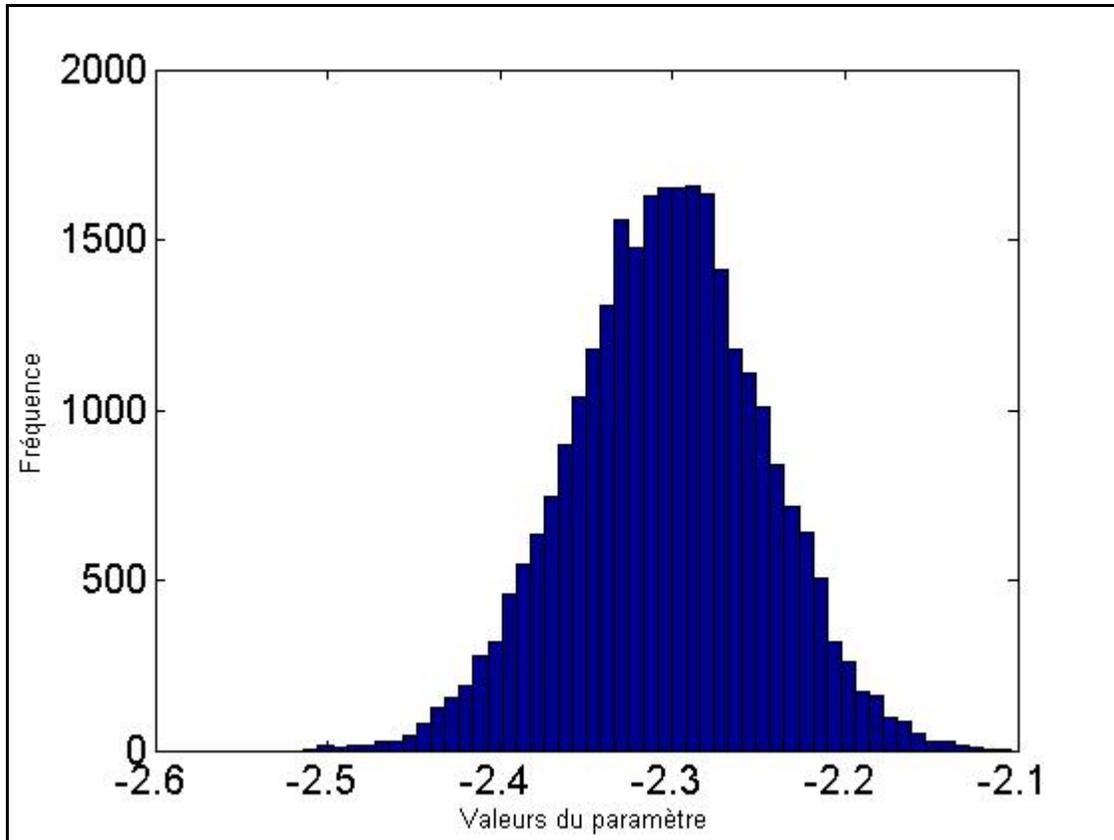
Les graphiques qui suivent présentent les distributions à posteriori des paramètres du modèle (5.16). Les traces des paramètres sont réduites aux 400 derniers tirages afin de présenter avec plus de précision les variations des paramètres lors de l'algorithme M-H.

Graphique 5 : Distribution à posteriori du paramètre  $\gamma_0$  dans le modèle (5.16) pour les 8 019 premières transactions avec changement de prix de l'action IBM



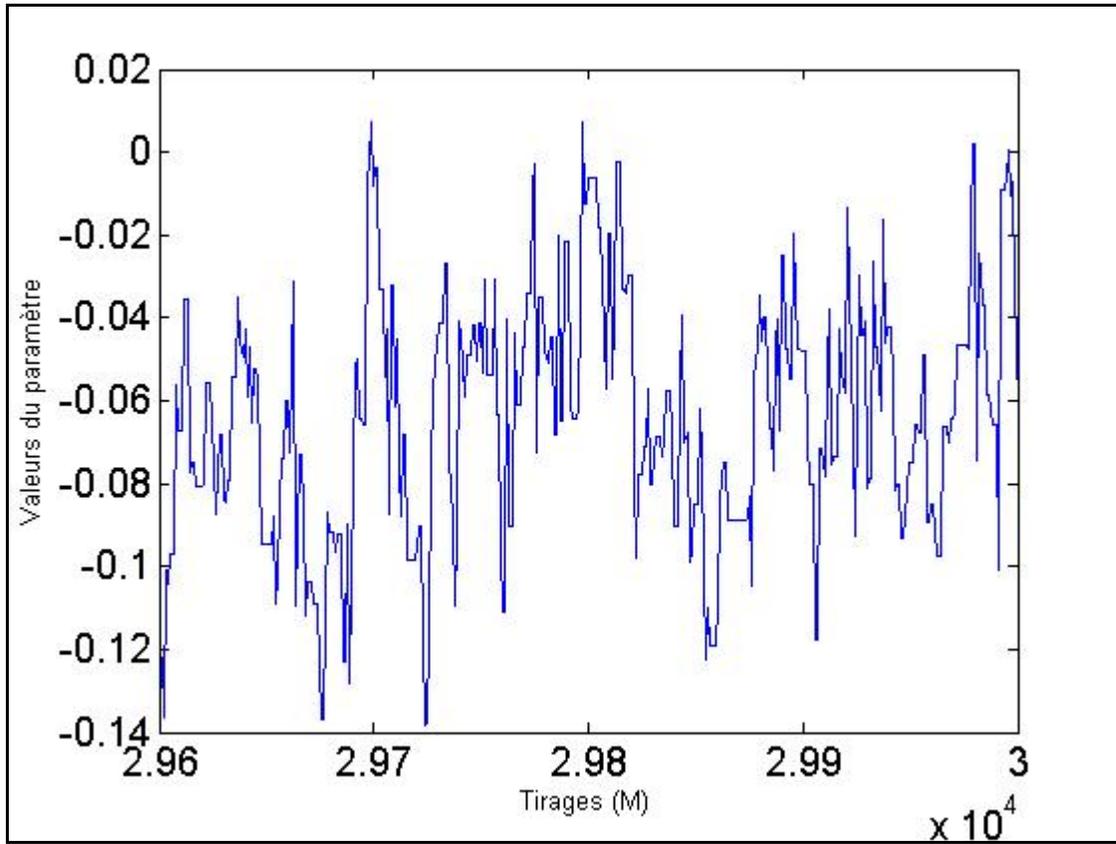
Source : Estimations du chercheur

Graphique 6 : Distribution à posteriori du paramètre  $\gamma_1$  dans le modèle (5.16) pour les 8 019 premières transactions avec changement de prix de l'action IBM



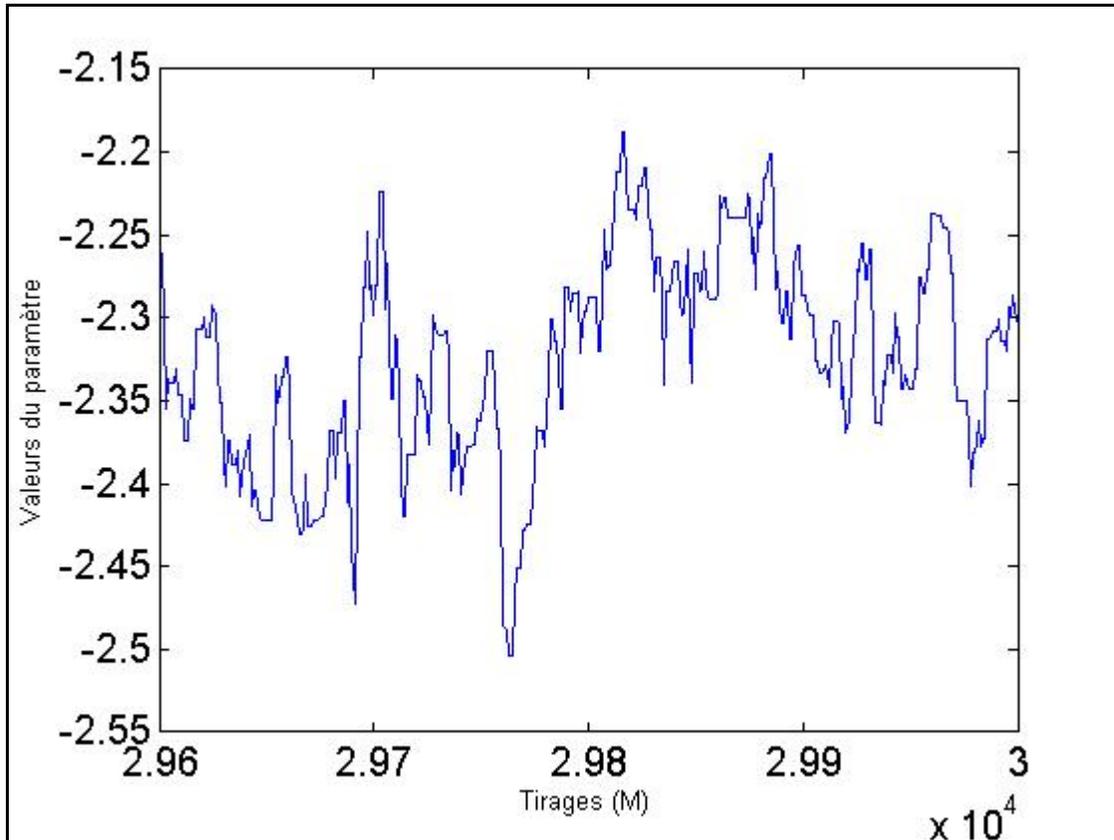
Source : Estimations du chercheur

Graphique 7 : Trace du paramètre  $\gamma_0$  dans le modèle (5.16) lors des 400 derniers tirages de l'algorithme M-H pour les 8 019 premières transactions avec changement de prix de l'action IBM



Source : Estimations du chercheur

Graphique 8 : Trace du paramètre  $\gamma_1$  dans le modèle (5.16) lors des 400 derniers tirages de l'algorithme M-H pour les 8 019 premières transactions avec changement de prix de l'action IBM



Source : Estimations du chercheur

La convergence des paramètres étant très rapide, il est intéressant d'observer plus en profondeur la fin de l'algorithme. Les graphiques 7 et 8 illustrent bien le degré d'autocorrélation qui prévaut. On remarque la présence de plateaux dans les deux traces, ce qui indique que la valeur proposée à ces tirages a été refusée car la valeur demeure la même qu'au tirage précédent.

### 5.3.3 Résultats pour la composante S

Pour la composante S, les modèles estimés successivement sont les suivants :

$$\ln\left(\frac{\lambda_{j,i}}{1-\lambda_{j,i}}\right) = \theta_{j,0} \quad (5.19)$$

$$\ln\left(\frac{\lambda_{j,i}}{1-\lambda_{j,i}}\right) = \theta_{j,0} + \theta_{j,1}S_{i-1} \quad (5.20)$$

$$\ln\left(\frac{\lambda_{j,i}}{1-\lambda_{j,i}}\right) = \theta_{j,0} + \theta_{j,1}S_{i-1} + \theta_{j,2}BA_i \quad (5.21)$$

$$\ln\left(\frac{\lambda_{j,i}}{1-\lambda_{j,i}}\right) = \theta_{j,0} + \theta_{j,1}S_{i-1} + \theta_{j,2}BA_i + \theta_{j,3}V_{i-1} \quad (5.22)$$

où  $j = u, d$  pour tous les modèles. Dans tous les modèles, l'ampleur des changements de prix est conditionnée sur la direction, soit positive (u) ou négative (d).

Pour l'étude de cette composante S (ampleur), on ne se sert que des données pour lesquelles un changement de prix est observé. Des 8 019 premières transactions enregistrées avec un changement de prix, on dénombre 4 023 hausses et 3996 baisses. Les résultats sont présentés dans les tableaux suivants :

*Tableau 10 : Estimation bayésienne des paramètres du modèle (5.19) pour les 4 023 premières transactions avec hausse de prix de l'action IBM*

Variable	$\bar{\theta}_u$	$\hat{\sigma}_{\theta_u}$	$\hat{S}_{\theta_u}$	$\hat{K}_{\theta_u}$	$\overline{\alpha(\theta_u^{(m-1)}, \theta_u^*)}$	$\overline{L(S_i^{(B)}   S_i^{(A)}, \theta_u^{(A)})}$
Cte	1.6523	0,0395	0,0340	2,9062	0,7517	-3 375,0

Source : Estimations du chercheur

Tableau 11 : Estimation bayésienne des paramètres du modèle (5.20) pour les 4 023 premières transactions avec hausse de prix de l'action IBM

Variable	$\bar{\theta}_u$	$\hat{\sigma}_{\theta_u}$	$\hat{S}_{\theta_u}$	$\hat{K}_{\theta_u}$	$\overline{\alpha(\theta_u^{(m-1)}, \theta_u^*)}$	$\overline{L(S_i^{(B)}   S_i^{(A)}, \theta_u^{(A)})}$
Cte	2,3356	0,0575	-0,0213	2,8925	0,5878	-3 141,9
$S_{i-1}$	-0,6917	0,0373	-0,0444	2,9008		

Source : Estimations du chercheur

Tableau 12 : Estimation bayésienne des paramètres du modèle (5.21) pour les 4 023 premières transactions avec hausse de prix de l'action IBM

Variable	$\bar{\theta}_u$	$\hat{\sigma}_{\theta_u}$	$\hat{S}_{\theta_u}$	$\hat{K}_{\theta_u}$	$\overline{\alpha(\theta_u^{(m-1)}, \theta_u^*)}$	$\overline{L(S_i^{(B)}   S_i^{(A)}, \theta_u^{(A)})}$
Cte	2,5754	0,0866	0,0620	3,2551	0,6759	-3 106,8
$S_{i-1}$	-0,6844	0,0383	-0,0516	2,9499		
$BA_i$	-0,5672	0,1532	0,0170	3,3009		

Source : Estimations du chercheur

Tableau 13 : Estimation bayésienne des paramètres du modèle (5.22) pour les 4 023 premières transactions avec hausse de prix de l'action IBM

Variable	$\bar{\theta}_u$	$\hat{\sigma}_{\theta_u}$	$\hat{S}_{\theta_u}$	$\hat{K}_{\theta_u}$	$\overline{\alpha(\theta_u^{(m-1)}, \theta_u^*)}$	$\overline{L(S_i^{(B)}   S_i^{(A)}, \theta_u^{(A)})}$
Cte	2,5741	0,0802	-0,0956	2,9138	0,3452	-3 107,3
$S_{i-1}$	-0,6841	0,0377	-0,1637	3,0294		
$BA_i$	-0,5512	0,1331	0,2610	3,3726		
$V_{i-1}$	-0,0047	0,0095	0,2330	3,2215		

Source : Estimations du chercheur

Tableau 14 : Estimation bayésienne des paramètres du modèle (5.19) pour les 3 996 premières transactions avec baisse de prix de l'action IBM

Variable	$\bar{\theta}_d$	$\hat{\sigma}_{\theta_d}$	$\hat{S}_{\theta_d}$	$\hat{K}_{\theta_d}$	$\overline{\alpha(\theta_d^{(m-1)}, \theta_d^*)}$	$\overline{L(S_i^{(B)}   S_i^{(A)}, \theta_d^{(A)})}$
Cte	1.6745	0,0405	-0,0373	3,1094	0,7578	-3 629,1

Source : Estimations du chercheur

Tableau 15 : Estimation bayésienne des paramètres du modèle (5.20) pour les 3 996 premières transactions avec baisse de prix de l'action IBM

Variable	$\bar{\theta}_d$	$\hat{\sigma}_{\theta_d}$	$\hat{S}_{\theta_d}$	$\hat{K}_{\theta_d}$	$\overline{\alpha(\theta_d^{(m-1)}, \theta_d^*)}$	$\overline{L(S_i^{(B)}   S_i^{(A)}, \theta_d^{(A)})}$
Cte	2,2262	0,0539	0,0920	2,8737	0,5622	-3 186,8
$S_{i-1}$	-0,4350	0,0299	-0,0661	3,0108		

Source : Estimations du chercheur

Tableau 16 : Estimation bayésienne des paramètres du modèle (5.21) pour les 3 996 premières transactions avec baisse de prix de l'action IBM

Variable	$\bar{\theta}_d$	$\hat{\sigma}_{\theta_d}$	$\hat{S}_{\theta_d}$	$\hat{K}_{\theta_d}$	$\overline{\alpha(\theta_d^{(m-1)}, \theta_d^*)}$	$\overline{L(S_i^{(B)}   S_i^{(A)}, \theta_d^{(A)})}$
Cte	2,3774	0,0791	0,1060	3,1187	0,6507	-3 176,6
$S_{i-1}$	-0,4353	0,0298	-0,0449	2,9049		
$BA_i$	-0,3532	0,1396	-0,0683	2,8784		

Source : Estimations du chercheur

Tableau 17 : Estimation bayésienne des paramètres du modèle (5.22) pour les 3 996 premières transactions avec baisse de prix de l'action IBM

Variable	$\bar{\theta}_d$	$\hat{\sigma}_{\theta_d}$	$\hat{S}_{\theta_d}$	$\hat{K}_{\theta_d}$	$\overline{\alpha(\theta_d^{(m-1)}, \theta_d^*)}$	$\overline{L(S_i^{(B)}   S_i^{(A)}, \theta_d^{(A)})}$
Cte	2,4232	0,0869	-0,0201	2,6273	0,2407	-3 175,9
$S_{i-1}$	-0,4354	0,0301	-0,0921	2,9988		
$BA_i$	-0,3756	0,1538	-0,0370	2,5746		
$V_{i-1}$	-0,0149	0,0063	0,1020	2,9661		

Source : Estimations du chercheur

Contrairement aux estimations faites pour les composantes A et D, la variable  $V_{i-1}$  n'améliore aucunement la qualité des prévisions de la composante S. On remarque même que la valeur moyenne des vraisemblances prédictives diminue dans le tableau 13. Cependant, la variable  $BA_i$  a beaucoup plus d'impact que dans les modèles des composantes A et D si l'on se fie aux moyennes des vraisemblances prédictives des tableaux 12 et 16. De plus, elle affecte considérablement la probabilité d'observer une hausse de prix de 1/8 de dollar ( $S_i = 1$ ) sachant l'ampleur du dernier changement et que la transaction implique une hausse du prix. Dans le contexte du MLE, on a les probabilités suivantes pour le modèle (5.3) :

$$P(S_i = 1 | S_{i-1} = 1, D_i = 1) = 0,827$$

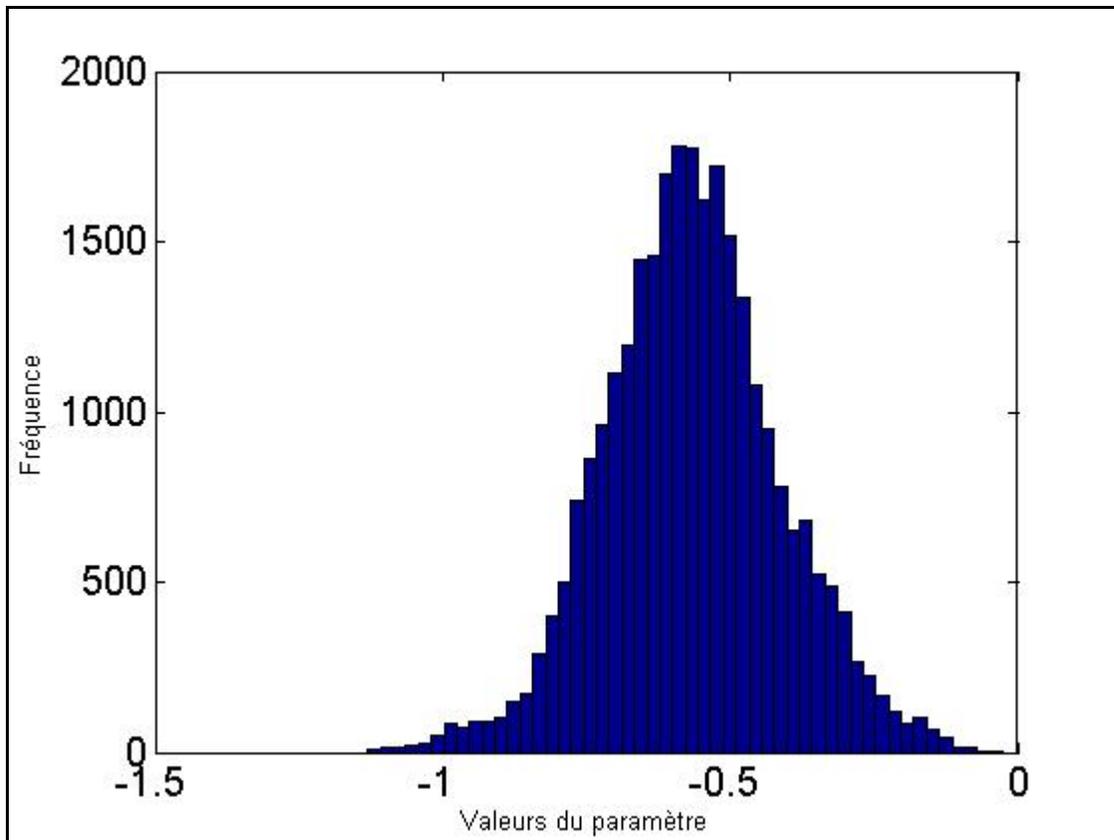
$$P(S_i = 1 | S_{i-1} = 2, D_i = 1) = 0,709$$

$$P(S_i = 1 | S_{i-1} = 3, D_i = 1) = 0,556$$

En conditionnant sur les valeurs moyennes des paramètres (voir le tableau 12), les probabilités calculées pour le modèle (5.21) sont respectivement de 0,861, 0,757, 0,611 si  $BA_i = 1/8$  \$ et de 0,833, 0,716, 0,559 si  $BA_i = 4/8$  \$. Ainsi, avec une grande valeur de  $BA_i$ , la probabilité d'observer un grand  $S_i$  semble également proportionnelle à  $S_{i-1}$  dans un contexte d'estimation bayésienne.

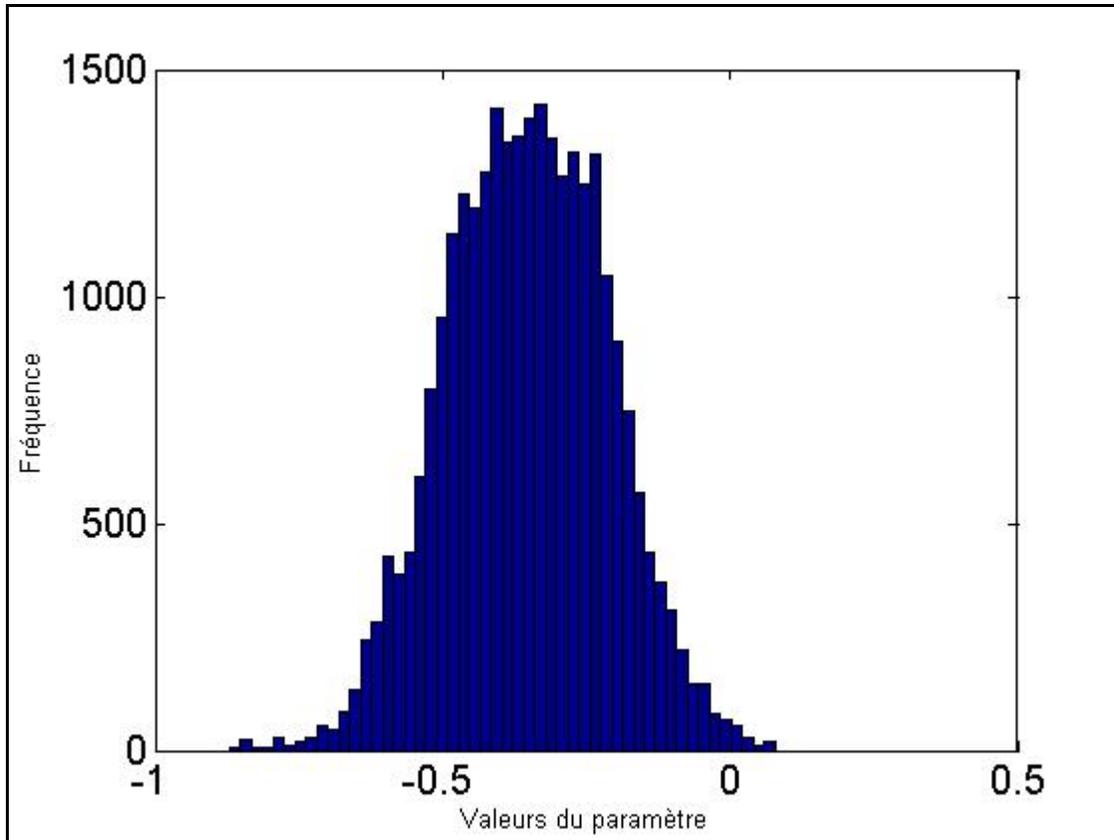
Les prochains graphiques présentent la distribution à posteriori et la trace des paramètres  $\theta_{u,2}$  et  $\theta_{d,2}$  associés à la variable  $BA_i$ . On s'intéresse plus particulièrement à ces paramètres puisqu'ils apportent beaucoup plus d'information pour la composante S que pour les composantes A et D.

*Graphique 9 : Distribution à posteriori du paramètre  $\theta_{u,2}$  dans le modèle (5.21) pour les 4 023 premières transactions avec hausse de prix de l'action IBM*



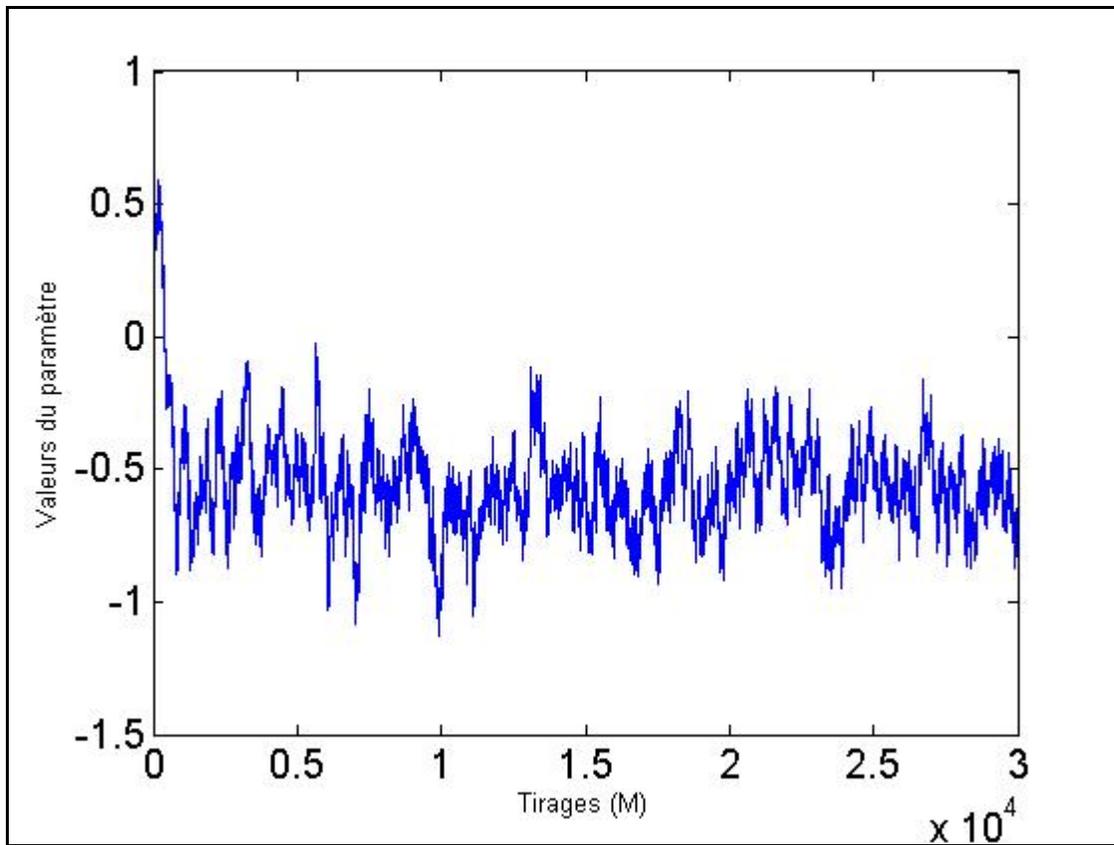
*Source : Estimations du chercheur*

Graphique 10 : Distribution à posteriori du paramètre  $\theta_{a,2}$  dans le modèle (5.21) pour les 3 996 premières transactions avec baisse de prix de l'action IBM



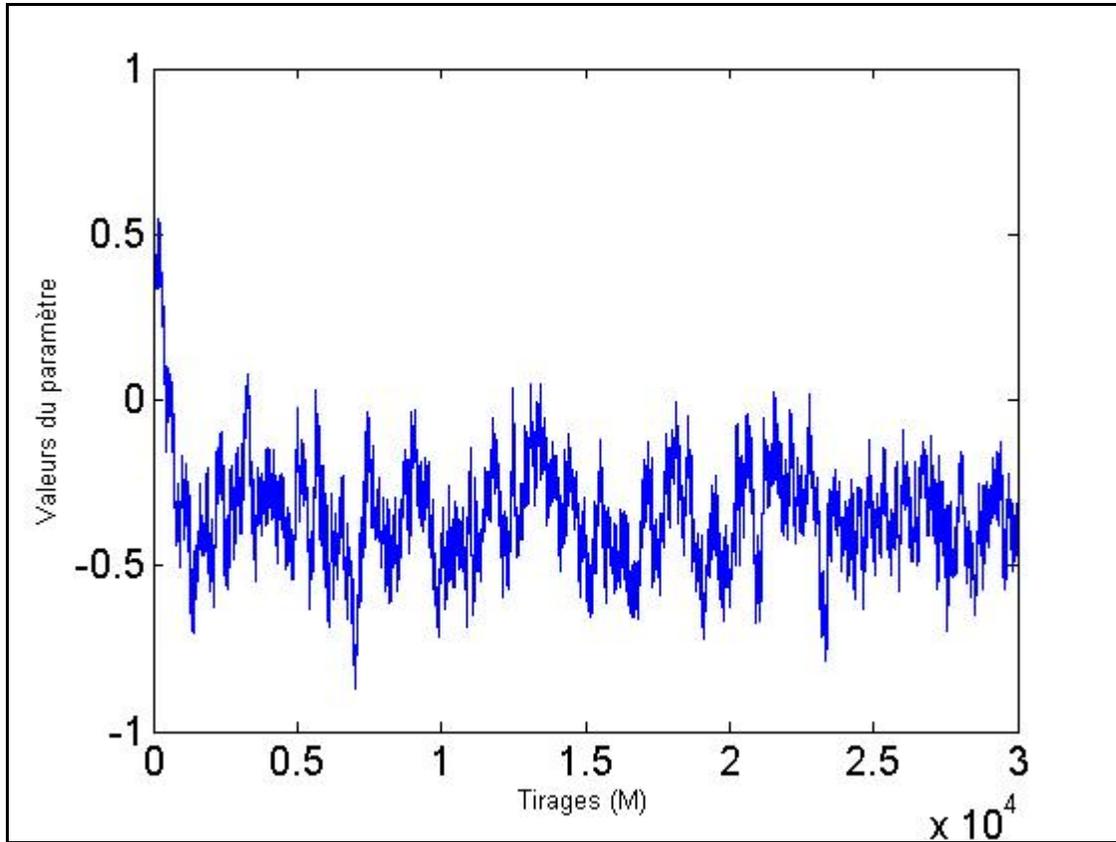
Source : Estimations du chercheur

Graphique 11 : Trace du paramètre  $\theta_{u,2}$  dans le modèle (5.21) lors de l'algorithme M-H pour les 4 023 premières transactions avec hausse de prix de l'action IBM



Source : Estimations du chercheur

Graphique 12 : Trace du paramètre  $\theta_{d,2}$  dans le modèle (5.21) lors de l'algorithme M-H pour les 3 996 premières transactions avec baisse de prix de l'action IBM



Source : Estimations du chercheur

En jetant un coup d'œil aux deux traces, on remarque que les valeurs des paramètres  $\theta_{u,2}$  et  $\theta_{d,2}$  sont plus volatiles que celles des autres. De plus, la convergence se fait moins rapidement. Il est toujours possible de réduire cette volatilité en augmentant le nombre de tirages dans l'algorithme M-H. Cependant, cette alternative étant très coûteuse en ressources, le chercheur a dû se contenter de ces résultats.

En rassemblant les résultats obtenus pour les trois composantes, on convient que la tentative d'ajouter des variables explicatives a plus ou moins réussi au niveau statistique car seulement la variable  $BA_t$  a un impact tangible sur les modèles étudiés. De plus, les valeurs des paramètres estimés par les deux approches sont très similaires dans tous les cas. Néanmoins, les résultats sont très intéressants sous la perspective économique. D'une part, ils suggèrent que la variable  $V_{i-1}$  est utile pour prédire la présence d'un changement de prix et sa direction. D'autre part, la variable  $BA_t$  améliore les prévisions pour l'ampleur de ces changements. Par ailleurs, les distributions à posteriori précises et la convergence rapide pour la plupart des paramètres estimés sont des indications de l'efficacité de l'algorithme M-H. Dans la prochaine section, une récapitulation des résultats obtenus par les deux méthodes est proposée.

## 6 Conclusion

Le but principal de cette étude était d'estimer avec la méthode bayésienne le modèle de décomposition des changements de prix ADS développé par Rydberg et Shephard (1998). Les objectifs rattachés à ce but étaient d'enrichir le modèle avec plus de variables explicatives et de comparer les résultats à ceux obtenus par la méthode d'estimation classique du maximum de vraisemblance.

Les données utilisées dans ce rapport sont les transactions de l'action IBM enregistrées au New York Stock Exchange (NYSE) entre 1<sup>er</sup> novembre 1990 et le 31 janvier 1991. On y dénombre 63 jours de transactions et 59 838 échanges de l'action.

Les informations pertinentes au préalable sont le caractère discret des données, les éléments de l'analyse bayésienne ainsi que la connaissance de l'outil de simulation Metropolis-Hastings.

L'analyse théorique du modèle ADS a consisté à dériver la fonction de vraisemblance et à déterminer une loi à priori. Ne sachant pas les valeurs possibles des paramètres à estimer à priori, le choix de la loi s'est arrêté sur la normale multivariée  $N(0, \Sigma)$ .

L'analyse empirique fait état des résultats de l'estimation du modèle ADS par le maximum de vraisemblance et par l'analyse bayésienne. Les résultats obtenus par les deux approches sont sensiblement les mêmes statistiquement parlant. Dans le cadre du maximum de vraisemblance, la seule variable explicative significative est le changement de prix à la transaction précédente. En adoptant la méthode bayésienne, la variable  $BA_i$  s'avère utile pour modéliser l'ampleur des changements. Elle vient confirmer la dépendance dynamique qui existe dans les changements de prix de titres échangés à haute fréquence. De plus, l'information sur les paramètres est plus complète dans un contexte bayésien car on peut caractériser leur distribution et leur incertitude. Même si l'impact statistique des variables ajoutées aux modèles estimés préalablement par l'approche classique n'est pas aussi déterminant que ne l'aurait souhaité le chercheur, les résultats ont une interprétation économique intéressante.

La modélisation des changements de titres boursiers est un sujet vaste et les pistes de recherches futures sont nombreuses. Dans un premier lieu, il serait intéressant de reprendre le modèle exploité dans cette étude avec des données d'autres titres et plus récentes. Le nombre de transactions de titres boursiers est de loin supérieur aujourd'hui qu'en 1991. L'un des objectifs du présent rapport étant de comparer des résultats obtenus selon deux méthodes d'estimation différentes, le choix du titre et de la période de transaction était une contrainte majeure. Deuxièmement, l'analyse bayésienne pourrait être appliquée à d'autres types de modèles afin de déterminer si cette approche peut être une valeur ajoutée aux méthodes classiques d'estimation dans la prévision des prix de titres boursiers. Finalement, il serait fort pertinent de déterminer si d'autres variables comme la durée entre les transactions ont réellement un pouvoir prédictif ou si les prix futurs d'un titre sont tout simplement trop volatiles.

## Bibliographie

Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton University Press: New Jersey.

Darsinos, T., and Satchell, S. (2001), "Bayesian Forecasting of Option Prices: A Natural Framework for Pooling Historical and Implied Volatility Information", Working paper, Department of Applied Economics, University of Cambridge.

Engle, R. F., and Russell, J. R. (1998), "Autoregressive conditional duration: A new model for irregularly spaced transaction data", *Econometrica*, 66, 1127-1162.

Hauseman, J., Lo, A., and MacKinlay, C. (1992), "An ordered probit analysis of transaction stock prices", *Journal of Financial Economics*, 31, 319-379.

Jacquier, E., Polson, N. G., and Rossi, P. E. (1994), "Bayesian analysis of stochastic volatility models", *Journal of Business & Economic Statistics*, 12, 371-417.

Koop, G. (2003), *Bayesian Econometrics*, John Wiley & Sons.

Li, K. (1999), "Bayesian Analysis of Duration Models: An Application to Chapter 11 Bankruptcy", *Economics Letters*, 63, 305-312.

McCulloch, R. E., and Tsay, R. S. (2001), "Nonlinearity in high frequency data and hierarchical models", *Studies in Nonlinear Dynamics & Econometrics*: Vol. 5: No. 1, Article 1.

Rydberg, T. H. and Shephard, N. (1998), "Dynamics of trade-by-trade price movements: decomposition and models", Working paper, Nuffield College, Oxford University.

Tsay, R. S. (2002), *Analysis of Financial Time Series*, Wiley Series in Probability and Statistics.

Wooldridge, J. M. (2003), *Introductory Econometrics: A Modern Approach*, 2e South-Western, Thomson Learning.