

A1.1

9

1022

Analyse Bayésienne des Modèles de Durées des Transactions Financières de Haute Fréquence

Par: Constant Aimé LONKENG NGOUANA

Dirigé par: William MCCAUSLAND

Département des Sciences Économiques, Université de Montréal

December 20, 2005

Contents

1	Résumé	2
2	Introduction	3
3	Revue des études antérieures	5
4	Informations pertinentes	8
4.1	Les données	8
4.2	L'approche bayésienne	8
5	Analyse Théorique	10
5.1	Le modèle ACD de base: "Autoregressive Conditional Du- ration model"	10
5.2	La vraisemblance du modèle	11
6	Analyse empirique	15
6.1	Une application de la méthode MCMC : l'algorithme Metropo- lis	15
6.2	Ajustement des données	17
6.3	Les résultats des estimations	20
6.3.1	Quelques graphiques des simulations	22
6.4	prévision du nombre de transactions au cours des T prochaines secondes	26
7	Conclusion	28

1 Résumé

Dans ce rapport de recherche, nous faisons une analyse bayésienne du modèle Conditionnel Autoregressif de Durées (ACD) introduit par Engle et Russell [7] en 1998 pour étudier la dynamique du temps qui s'écoule entre les transactions successives d'une action en bourse. Nous utilisons la méthode Markov Chain Monte Carlo (MCMC) et plus précisément l'algorithme Metropolis. Il nous permet de simuler la distribution à posteriori des paramètres par un principe qui consiste à parcourir l'espace des paramètres en sélectionnant chaque fois un candidat dont la vraisemblance rapportée à celle du candidat précédemment retenu excède un seuil de tolérance donné par un tirage de la loi uniforme. La base de données utilisée pour cette étude regroupe toutes les transactions ayant eu cours sur l'action IBM à la "New York Stock Exchange" (NYSE) entre le 1er novembre 1990 et le 31 janvier 1991. Les moyennes à posteriori des paramètres sont assez proches des estimations ponctuelles obtenues dans ce même rapport par application de la méthode classique du Quasi-Maximum de Vraisemblance. Il en ressort que l'espérance de durée conditionnelle est très persistente et que la spécification Weibull serait plus adéquate que la spécification exponentielle pour la distribution des innovations. Notre approche a l'avantage de présenter des résultats valides en échantillon fini. Nous l'utilisons pour prédire le nombre de transactions susceptibles d'être transigées en bourse dans un intervalle de temps donné.

2 Introduction

Au cours des sept dernières années, une littérature croissante s'est développée autour de la modélisation du temps qui s'écoule entre deux transactions successives d'une action en bourse. Une telle expansion a été favorisée par le développement rapide de la technologie qui nous permet aujourd'hui de stocker les informations massives (temps, volume, prix, bid, spread, ask) à la seconde près sur les titres transigés en bourse.

L'intérêt d'étudier le temps qui s'écoule entre les transactions d'une action est de comprendre la manière dont l'information pénètre le marché financier comme le mentionnent Easley et O'Hara [5]. Par exemple, des courtes durées entre les transactions matérialisent l'entrée rapide d'une nouvelle information sur le marché et réciproquement.

Cependant, les données de haute fréquence ont la particularité d'être irrégulièrement espacées du fait que le temps d'occurrence des transactions possède une composante aléatoire importante. Ceci rend inappropriées certaines techniques de l'économétrie standard pour laquelle les variables sont observées à une fréquence bien déterminée. On parle par exemple de PIB mensuel, trimestriel, annuel. Une manière de traiter des données espacées de façon aléatoire est leur agrégation en intervalles régulièrement espacés. Cette approche, bien que simple et pratique présente l'inconvénient de dissiper certaines informations contenues dans les micro-données. Par exemple, avoir des informations sur la volatilité d'un actif sur une base journalière ne nous permet pas de cerner la variabilité du rendement de cet actif à différentes périodes de la journée.

Dans leur travail séminal, Engle et Russell [7] proposent pour les données de haute fréquence un modèle de Durées Autoregressif Conditionnel (ACD) qui est en fait un ARMA avec des innovations non-gaussiennes. Ce modèle est pour les durées ce qu'est pour la volatilité le célèbre modèle ARCH développé par Engle en 1982 [6] pour prendre en compte la non constance de la variance des rendements d'actifs financiers.

Malgré le fait que le modèle ACD ait connu de nombreuses extensions depuis son introduction en 1998, la littérature reste très peu féconde quant à l'application des méthodes bayésiennes à cette analyse. Dans ce rapport, nous essayons de surmonter ce manquement. Il s'agit essentiellement d'utiliser l'algorithme Metropolis de la classe MCMC (Markov Chain Monte Carlo) pour simuler les densités à posteriori des paramètres du modèle ACD. L'approche bayésienne nous semble assez intuitive car elle traduit la réalité selon laquelle l'économètre part généralement avec

une idée assez vague des paramètres (loi à priori), idées qui s'éclaircissent au fur et à mesure qu'il observe le phénomène modélisé. Ainsi, nous avons une précision de plus en plus grande des paramètres à la lumière des données.

Notre travail est organisé comme suit: Après cette section introductive, nous faisons un survol de la littérature développée jusqu'à date sur les modèles de durées des transactions financières de haute fréquence. Il est ensuite question dans la troisième section de présenter quelques informations pertinentes, notamment les données et une brève introduction à l'approche bayésienne de l'économétrie. Dans la section cinq consacrée à l'analyse théorique du modèle ACD de base, nous en dérivons la vraisemblance ainsi que les lois à posteriori de nos paramètres. L'analyse empirique fait l'objet de la section six et exploite l'un des mérites de l'approche bayésienne par un algorithme de prévision du nombre de transactions dans un intervalle de temps donné. Nous concluons enfin dans la dernière section.

3 Revue des études antérieures

Le travail séminal de Engle et Russell [7] a ouvert la voie à une immense littérature sur les modèles de durées. Cette littérature est rendue féconde par les capacités de stockage croissantes et la possibilité technologique d'enregistrer toutes les transactions ayant cours sur un titre transigé en bourse. Le fait que les transactions surviennent de façon aléatoire imprègne une structure particulière aux données de haute fréquence qui sont alors espacées de façon très irrégulière. Ceci est une véritable démarcation par rapport aux séries temporelles classiques pour lesquelles les données sont espacées par un intervalle de temps de même longueur (jour, mois, année, etc.). Une manière de traiter des données ainsi espacées de façon aléatoire est de les agréger en intervalles de même longueur. Mais alors il se pose un dilemme. Si on choisit des intervalles de temps trop longs, on perd certaines informations pertinentes qui pourraient être contenues dans les micro-données. Par contre si on choisit des intervalles de temps trop courts on court également le risque d'avoir plusieurs intervalles sans nouvelle information ce qui introduit une forme particulière d'hétéroscédasticité dans les données, comme le mentionnent Engle et Russell [7]. Ils proposent, dans l'esprit du ARCH un modèle qui prend en compte cette forme particulière des données en modélisant la durée entre les transactions comme étant un processus autoregressif.

Dès lors, beaucoup de variantes de ce modèle de base ont été proposées. Bauwens et Giot [2] en introduisent la forme logarithmique qui permet l'existence d'une relation non linéaire entre les durées et leurs valeurs retardées. Ce modèle "log-ACD" est également intéressant en ce sens qu'il permet d'éliminer la contrainte de non-négativité des coefficients. En effet, même avec un logarithme négatif, la positivité de la durée conditionnelle (obtenue par passage à l'exponentielle) reste garantie. Cette propriété se révèle particulièrement pertinente lorsqu'on veut introduire des éléments autres que les variables retardées des durées et de leur espérance conditionnelle dans l'analyse.

D'autres pistes de recherche sont liées à la distribution des innovations. Ceci s'avère intéressant en ce sens que les innovations matérialisent la différence qui existerait entre la durée conditionnelle à estimer et la durée réellement observée. Dans le modèle de base elles suivent une loi exponentielle ou Weibull qui sont des fonctions hasard assez utilisées dans la littérature plus générale des modèles de durées dans un état

donné. Greene [13] en fait une description détaillée. Grammig et Maurer [9] considèrent des innovations suivant une distribution Burr précédé par Lunde [17] qui introduit des innovations de loi gamma généralisée. Ces deux dernières formes d'innovation présentent l'avantage d'englober les distributions exponentielle et Weibull comme cas particuliers.

Bauwens et Veredas [1] introduisent le modèle de Durées Conditionnel Stochastique (SCD). L'idée ici est que les durées entre les transactions sont dictées par une variable aléatoire latente (non observée) et dynamique. Ceci est l'équivalent du modèle SV développé pour la volatilité des actifs financiers par Melino et Turnbull [18]. L'estimation du modèle peut se faire par la méthode de quasi-maximum de vraisemblance en utilisant le filtre de Kalman. Cependant, comme le remarque Tsay [20], il est difficile à estimer du fait de la présence de 2 innovations. Jacquier, Polson et Rossi [14] comparent cette méthode d'estimation à celle de Markov Chain Monte Carlo dans le cas du modèle SV.

Ghysels, Gouriéroux et Jasiak [12], dans un esprit prolongé du SCD introduisent le modèle de durées à volatilité stochastique (SVD). Contrairement au modèle SCD qui est à un seul facteur, le SVD est un modèle à deux facteurs qui prend en compte non seulement la dynamique de la durée mais aussi la dynamique de sa variance (volatilité).

Contrairement à toutes les études sus-mentionnées, Ola Simonsen [19] suggère une approche multivariée pour prendre en compte le fait qu'une information nouvelle affecte non pas un seul titre transigé en bourse mais est susceptible d'en affecter plusieurs simultanément. Cette approche multivariée offre la possibilité d'une analyse de causalité entre les délais de transactions de différentes actions. Il utilise ce modèle pour étudier les durées entre les transactions de deux actions transigées à la bourse des valeurs de Stockholm : l'action Ericsson B et l'action AstraZeneca. Il en ressort que l'ajout d'une deuxième action augmente effectivement le pouvoir explicatif du modèle univarié de durées. L'analyse de Causalité dans ce cas révèle que Ericsson B Granger cause les durées de AstraZeneca, tandis que la réciproque n'est pas vraie. Ce résultat n'est guère surprenant au regard de la taille du titre Ericsson B comparée à celle du titre AstraZeneca sur la bourse de Stockholm.

Cependant, ces études antérieures, bien que très denses, reflet de l'intérêt croissant pour un sujet très récent présentent la principale faiblesse de se fonder exclusivement sur l'approche classique de l'économétrie. Dans un contexte de prévision, cette approche ne prend pas en compte l'incertitude qui plane sur les paramètres du modèle. L'approche bayésienne consiste à partir d'une loi à priori assez imprécise des paramètres

pour aboutir à une loi à posteriori plus précise à la lumière des données. Aussi, les prévisions sont faites en termes de densité prédictive pour prendre en compte l'incertitude qui entoure les paramètres.

Pour rendre le texte plus digeste, nous présentons tout d'abord quelques informations pertinentes pour notre étude telles que les données et une introduction à l'approche bayésienne.

4 Informations pertinentes

4.1 Les données

Les données utilisées dans ce rapport regroupent toutes les transactions ayant eu cours sur l'action IBM entre le 1er novembre 1990 et le 31 janvier 1991 à la New York Stock Exchange (NYSE). Cette base de données, disponible sur les pages web de Tsay et Engle donne pour chaque transaction la seconde à laquelle elle a eu lieu, le prix, le volume, le bid, le spread et le ask. Seul le temps auquel a eu lieu la transaction nous intéresse dans la présente étude. Notons t_i le temps auquel s'effectue la transaction i . La durée d_i mesure le temps qui s'écoule entre la transaction $i - 1$ et la transaction i et est donc donnée par $d_i = t_i - t_{i-1}$. Ces durées, comme nous l'expliquons dans la section empirique sont ensuite ajustées pour prendre en compte l'effet moment de la journée. Le modèle peut donc ensuite être estimé en utilisant les durées ajustées $x_i = \frac{\Delta t_i}{g(t_i)} = \frac{d_i}{g(t_i)}$ où g est une fonction déterministe devant être estimée.

Le même travail pourra être répliqué pour n'importe quelle action sur n'importe quel marché financier pour autant que les données de haute fréquence soient disponibles.

4.2 L'approche bayésienne

Dans l'approche classique de l'économétrie, les paramètres à estimer, bien que inconnus sont supposés fixes. Dans la littérature relative à ce travail, on procède soit par la méthode de Maximum de vraisemblance (ML) pour les innovations gaussiennes soit par Quasi-Maximum de vraisemblance (QML) dans le cas des innovations non-gaussiennes. L'idée dans les deux cas est de déterminer les paramètres qui cadrent le mieux avec les données observées, étant donnée la fonction de vraisemblance. On obtient alors des valeurs pour différents paramètres et on peut directement procéder aux prévisions sans tenir compte de l'incertitude qui existe dans les paramètres ainsi estimés.

Dans l'approche bayésienne par contre, l'économètre part d'une loi à priori des paramètres et obtient une loi à posteriori plus précise à la lumière des données. Les prévisions sont alors basées sur cette loi à posteriori, prenant ainsi en compte l'incertitude qui plane sur les paramètres. Cette approche nous semble assez naturelle dans la mesure où elle reflète l'idée d'apprentissage constatée dans la vie quotidienne.

En effet, l'agent économique part toujours avec une idée d'un phénomène et révisé ses anticipations au fur et à mesure qu'il observe les tendances du phénomène.

L'approche bayésienne comme son nom l'indique est fondée sur la règle de Bayes.

Notons θ le vecteur $k \times 1$ des paramètres inconnus du modèle et X le vecteur des données. D'après la règle de Bayes:

$$f(\theta | X) = \frac{f(\theta, X)}{f(X)} = \frac{f(X | \theta)p(\theta)}{f(X)} \quad (1)$$

où $p(\theta)$ représente la distribution à priori des paramètres,

$f(X | \theta)$ la fonction de vraisemblance au vue des données,

$f(X)$ la distribution marginale de X qui peut être obtenue par intégration sur l'ensemble des valeurs possibles de θ :

$$f(X) = \int_{\theta \in \Theta} f(X, \theta) d\theta = \int_{\theta \in \Theta} f(X | \theta)p(\theta) d\theta \quad (2)$$

$f(\theta | X)$ est appelée distribution à posteriori de θ . Lorsqu'on s'intéresse à la loi de $\theta | X$, $f(X)$ devient une constante dans l'équation (1) vu que la loi est conditionnelle à X . Il vient donc que:

$$f(\theta | X) \propto f(X | \theta)p(\theta) \quad (3)$$

La relation (3) est fondamentale pour l'analyse bayésienne¹ car elle nous permet d'avoir la loi à posteriori des paramètres à un facteur multiplicatif "constant" près. On peut alors faire appel aux méthodes MCMC pour tirer un échantillon de cette loi à posteriori et examiner le comportement complet des paramètres.

¹Nous y reviendrons plus tard dans la partie théorique de ce travail

5 Analyse Théorique

5.1 Le modèle ACD de base: "Autoregressive Conditional Duration model"

Soit x_i la durée ajustée² qui s'écoule entre la transaction $i - 1$ et la transaction i .

Notons:

$$\eta_i = E(x_i | I_{i-1}) \quad (4)$$

l'espérance conditionnelle de cette durée, où I_{i-1} désigne l'ensemble des informations disponibles avant la i ème transaction.

Le modèle $ACD(r, s)$ pour "Autoregressive Conditional Duration model" se définit comme suit :

$$x_i = \eta_i \epsilon_i \quad (5)$$

où (ϵ_i) est une séquence de variables aléatoires indépendantes et identiquement distribuées telle que $E(\epsilon_i) = 1$.

Dans le modèle de base, l'espérance conditionnelle de durée dépend linéairement de ses valeurs retardées et des valeurs retardées des durées:

$$\eta_i = \omega + \sum_{k=1}^r \gamma_k x_{i-k} + \sum_{k=1}^s \lambda_k \eta_{i-k} \quad (6)$$

Pour ce qui est des valeurs de r et s on se limite généralement au $ACD(1, 1)$ car les applications montrent que l'ajout des variables retardées rend le modèle moins parcimonieux sans contribution notable à son pouvoir explicatif. Le modèle retenu ainsi à

$$x_i = \eta_i \epsilon_i$$

$$\eta_i = \omega + \gamma x_{i-1} + \lambda \eta_{i-1} \quad (7)$$

où³

²c'est à dire la durée modifiée de façon à prendre en compte l'effet déterministique imprégné par le moment spécifique de la journée auquel la transaction a eu lieu.

³nous expliquons l'introduction de ce facteur multiplicatif des innovations par la suite.

$$\epsilon_i = \frac{\zeta_i}{\beta\Gamma\left(1 + \frac{1}{\alpha}\right)} \quad (8)$$

Nous considérons ici le cas où ζ_i suit une loi Weibull (on parle alors de WACD)⁴ de paramètres α, β ($\alpha > 0, \beta > 0$), dont la densité est:

$$h(\zeta_i | \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta^\alpha} \zeta_i^{\alpha-1} e^{-\left(\frac{\zeta_i}{\beta}\right)^\alpha} & \text{si } \zeta_i \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (9)$$

L'avantage de cette loi Weibull que nous utilisons pour nos estimations est qu'elle admet la loi exponentielle (qui a la caractéristique d'être sans mémoire) comme cas particulier. Il suffit pour cela de prendre $\alpha = 1$.

En introduisant le facteur $\frac{1}{\beta\Gamma\left(1 + \frac{1}{\alpha}\right)}$, ϵ_i suit une weibull standard.

En effet, on a:

$$E(\zeta_i) = \beta\Gamma\left(1 + \frac{1}{\alpha}\right) \quad (10)$$

$$\text{d'où } E(\epsilon_i) = E\left(\frac{\zeta_i}{\beta\Gamma\left(1 + \frac{1}{\alpha}\right)}\right) = \frac{1}{\beta\Gamma\left(1 + \frac{1}{\alpha}\right)} E(\zeta_i) = 1$$

Nous disposons donc de tous les éléments préalables pour déterminer la vraisemblance du modèle.

5.2 La vraisemblance du modèle

La vraisemblance du modèle pour T observations s'écrit :

$$\begin{aligned} f(x_0, x_1, x_2, \dots, x_T | \theta) &= f(x_1, x_2, \dots, x_T | \theta, x_0) \times f(x_0 | \theta) \\ &= f(x_1 | \theta, x_0) \times f(x_2 | \theta, x_0, x_1) \times \dots \\ &\quad \times f(x_T | \theta, x_0, x_1, x_2, \dots, x_{T-1}) \times f(x_0 | \theta) \end{aligned}$$

où θ désigne le vecteur des paramètres du modèle

⁴Dans leur travail séminal, Engle & Russell [7] considèrent successivement les cas où les innovations ϵ_i suivent une loi Exponentielle (EACD) et Weibull standardisée (WACD).

Posons: $X = (x_1, x_2, \dots, x_T)$. L'expression précédente peut alors se réduire à:

$$f(X | \theta, x_0) = \frac{f(x_0, X | \theta)}{f(x_0 | \theta)} = \prod_{i=1}^T f(x_i | I_{i-1}, \theta) \quad (11)$$

avec $f(x_i | I_{i-1}, \theta) = f(\eta_i \epsilon_i | \theta)$ où $\eta_i = \omega + \gamma x_{i-1} + \lambda \eta_{i-1}$ est connue à i .

En appliquant le Jacobien de transformation à l'expression $x_i = \eta_i \epsilon_i$ on obtient

$$\begin{aligned} f(x_i | I_{i-1}, \theta) &= f(\eta_i \epsilon_i | I_{i-1}, \theta)^5 = \frac{1}{\eta_i} h(\epsilon_i) \\ &= \frac{1}{\eta_i} h\left(\frac{\zeta_i}{\beta \Gamma(1 + \frac{1}{\alpha})}\right) \text{ d'après (8)} \\ &= \frac{1}{\eta_i} \beta \Gamma\left(1 + \frac{1}{\alpha}\right) h(\zeta_i) \\ &= \frac{1}{\eta_i} \beta \Gamma\left(1 + \frac{1}{\alpha}\right) h\left(\frac{x_i}{\eta_i} \beta \Gamma\left(1 + \frac{1}{\alpha}\right)\right) \text{ car } \zeta_i = \beta \Gamma\left(1 + \frac{1}{\alpha}\right) \times \epsilon_i \text{ et } \epsilon_i = \frac{x_i}{\eta_i} \end{aligned}$$

En remplaçant h par son expression donnée par (9) on obtient alors:

$$\begin{aligned} f(x_i | I_{i-1}, \theta) &= \frac{1}{\eta_i} \beta \Gamma\left(1 + \frac{1}{\alpha}\right) \frac{\alpha}{\beta^\alpha} \left[\frac{x_i}{\eta_i} \beta \Gamma\left(1 + \frac{1}{\alpha}\right)\right]^{\alpha-1} \exp\left(-\frac{1}{\beta} \frac{x_i}{\eta_i} \beta \Gamma\left(1 + \frac{1}{\alpha}\right)\right) \\ &= \frac{1}{\eta_i} \frac{\alpha}{\beta^{\alpha-1}} \left(\frac{x_i}{\eta_i}\right)^{\alpha-1} \beta^{\alpha-1} \Gamma\left(1 + \frac{1}{\alpha}\right) [\Gamma\left(1 + \frac{1}{\alpha}\right)]^{\alpha-1} \exp\left(\left[-\frac{x_i}{\eta_i} \Gamma\left(1 + \frac{1}{\alpha}\right)\right]^\alpha\right) \\ &= \frac{1}{\eta_i} \alpha \left(\frac{x_i}{\eta_i}\right)^{\alpha-1} [\Gamma\left(1 + \frac{1}{\alpha}\right)]^\alpha \exp\left(\left[-\frac{x_i}{\eta_i} \Gamma\left(1 + \frac{1}{\alpha}\right)\right]^\alpha\right) \\ &= \frac{1}{x_i} \alpha \left(\frac{x_i}{\eta_i}\right)^\alpha [\Gamma\left(1 + \frac{1}{\alpha}\right)]^\alpha \exp\left(\left[-\frac{x_i}{\eta_i} \Gamma\left(1 + \frac{1}{\alpha}\right)\right]^\alpha\right) \end{aligned}$$

On a donc d'après (11) l'expression finale ci-dessous de la vraisemblance⁶:

$$f(X | \omega, \gamma, \lambda, \alpha, x_0) = \prod_{i=1}^T \left\{ \frac{1}{x_i} \alpha \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^\alpha \left(\frac{x_i}{\eta_i}\right)^\alpha \times \exp\left[-\Gamma\left(1 + \frac{1}{\alpha}\right) \frac{x_i}{\eta_i}\right]^\alpha \right\} \quad (12)$$

⁵ car $x_i = \eta_i \epsilon_i$

⁶ notons que le terme β n'apparaît pas dans la vraisemblance dû au facteur $\frac{1}{\beta \Gamma(1 + \frac{1}{\alpha})}$ appliqué à l'innovation Weibull de paramètre (α, β) pour que son espérance soit égale à 1 comme motivé précédemment.

D'où la log-vraisemblance:

$$LnL(X | \omega, \gamma, \lambda, \alpha, x_0) = \sum_{i=1}^T \left\{ \begin{array}{l} \log\left(\frac{\alpha}{x_i}\right) + \alpha \log\left(\Gamma\left(1 + \frac{1}{\alpha}\right)\right) + \alpha \log\left(\frac{x_i}{\eta_i}\right) \\ - \left[\Gamma\left(1 + \frac{1}{\alpha}\right) \frac{x_i}{\eta_i}\right]^\alpha + \alpha \log\left(\frac{x_i}{\eta_i}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right) \frac{x_i}{\eta_i}\right]^\alpha \end{array} \right\} \quad (13)$$

Nous allons la maximiser dans la partie empirique pour trouver les valeurs des paramètres comme cela se fait dans l'approche classique avant de passer à l'analyse bayésienne proprement dite.

La loi à posteriori du vecteur des paramètres est obtenue à l'aide de la relation(3) explicitée à la section précédente.

La forme de la vraisemblance ci-haut n'admet pas de façon triviale une loi à priori conjuguée. En effet, il ne nous paraît pas possible de la factoriser en 2 termes: un premier terme, indépendant des observations et correspondant à la densité à priori et un second terme fonction des observations et des paramètres⁷. Faute d'informations supplémentaires, nous allons donc prendre la loi normale multivariée comme loi à priori de nos paramètres. Supposons donc:

$$\theta \sim N(\mu, \Omega) \quad (14)$$

où nous admettons des corrélations entre les paramètres à travers la matrice des variance-covariances Ω . En effet, il est tout à fait naturel de penser que la persistance existant dans l'espérance conditionnelle (mesurée par le paramètre λ) serait fonction de la façon dont elle est corrélée aux durées précédemment observées (paramètre γ). Par exemple, l'influence de l'espérance conditionnelle de la période précédente sera d'autant plus faible que l'influence de la durée de la période précédente est grande et réciproquement. L'idée ici est que la durée antérieurement observée capte un bout de ce qui aurait pu être expliqué par l'espérance conditionnelle si elle était considérée toute seule. Cependant, il n'est pas très naturel de penser que le paramètre α de la loi Weibull soit directement corrélé aux paramètres de persistance des durées que sont γ et λ . La dernière colonne de Ω , hors mis le dernier terme correspondant à la variance de α sera alors considéré comme nul.

Par définition de la loi normale multivariée on a:

⁷On permet également un facteur multiplicatif dépendant uniquement des observations.

$$p(\theta) = \frac{1}{(2\pi)^{k/2} |\Omega|^{1/2}} \exp\left(-\frac{1}{2}(\theta - \mu)' \Omega^{-1}(\theta - \mu)\right) \quad (15)$$

$$\propto \exp\left(-\frac{1}{2}(\theta - \mu)' \Omega^{-1}(\theta - \mu)\right)$$

Il vient donc que:

$$f(\theta | X) = \frac{1}{f(X)} \times \left\{ \prod_{i=1}^T \frac{\alpha}{x_i} \left[\Gamma\left(1 + \frac{1}{\alpha}\right) \right]^\alpha \left(\frac{x_i}{\eta_i} \right)^\alpha \exp\left[-\Gamma\left(1 + \frac{1}{\alpha}\right) \frac{x_i}{\eta_i}\right]^\alpha \right\}$$

$$\times \left\{ \exp\left(-\frac{1}{2}(\theta - \mu)' \Omega^{-1}(\theta - \mu)\right) \right\}$$

d'après (12) combinée avec (15) et en référence à la règle de Bayes donnée par (1)

$$\Rightarrow f(\theta | X) \propto \left\{ \prod_{i=1}^T \frac{\alpha}{x_i} \left(\frac{x_i}{\eta_i} \right)^\alpha \right\} \times \exp \left\{ \begin{array}{l} (-\Gamma(1 + \frac{1}{\alpha}))^\alpha \sum_{i=1}^T \left(\frac{x_i}{\eta_i} \right)^\alpha \\ -\frac{1}{2}(\theta - \mu)' \Omega^{-1}(\theta - \mu) \end{array} \right\}$$

par application de (3)

Cette loi est simulée dans la partie empirique suivante en faisant recours à l'algorithme-Métropolis.

6 Analyse empirique

6.1 Une application de la méthode MCMC : l'algorithme Metropolis

Il convient de noter que certains problèmes auxquels nous nous intéressons dans cette partie empirique consistent à évaluer des expressions de la forme $E[h(\theta) | X]$ où h est une fonction quelconque. Par exemple l'espérance à posteriori des paramètres est obtenue pour $h(\theta) = \theta$, ce qui correspond à une estimation ponctuelle. Le problème qui se pose est qu'il nous est impossible, ou tout au moins très difficile d'obtenir la solution d'un tel problème de façon analytique. Ceci est dû au fait que notre densité à posteriori et la densité à priori ne sont pas conjuguées (conjugate priors). On a alors recours aux simulations. Si nous pouvions simuler $\theta^{(m)}$ i.i.d de même loi $\theta | X$ alors le problème serait résolu par application du résultat asymptotique suivant :

$$\frac{1}{M} \sum_{m=1}^M h(\theta^{(m)}) \xrightarrow{ps} E[h(\theta) | X] \quad \text{lorsque } M \rightarrow \infty \quad (16)$$

Mais alors il est très souvent impossible d'obtenir une telle simulation de façon directe. C'est à ce niveau qu'interviennent les méthodes de simulation Markov Chain⁸ Monte Carlo (MCMC).

Nous construisons un processus markovien $\theta^{(m)} | \theta^{(m-1)}$ pour lequel $\theta | X$ est la loi stationnaire unique. Avec quelques conditions techniques minimales, le résultat (16) continue à tenir. La question est donc de savoir comment construire une telle chaîne. Plusieurs algorithmes ont ainsi été développés dans la littérature dans ce sens. Le plus utilisé est sans doute l'échantillonnage de Gibbs. Il est intéressant lorsque la distribution à posteriori jointe des paramètres est difficile à obtenir tandis que la distribution à posteriori de chaque paramètre, conditionnellement aux autres paramètres est facile à obtenir. Le principe consiste donc à

⁸Notons qu'une séquence $\theta^{(m)}$ est une chaîne de Markov si sa distribution conditionnelle satisfait la propriété suivante:

$$P(\theta^{(h)} | \theta^{(s)}, s \leq t) = P(\theta^{(h)} | \theta^{(t)}) \quad \forall h > t.$$

Cette propriété nous dit tout simplement qu'une chaîne de Markov est une séquence $\theta^{(m)}$ telle que étant donnée la valeur de $\theta^{(t)}$, les valeurs de $\theta^{(h)}$, $h > t$ ne dépendent pas des valeurs de $\theta^{(s)}$, $s < t$.

transformer le problème multidimensionnel complexe en une séquence de problèmes unidimensionnels plus faciles à résoudre. Mais dans notre cas, même les distributions à posteriori des paramètres pris isolément sont difficiles à simuler directement. Il existe pour ce genre de situations un algorithme alternatif : l'algorithme Metropolis (1949) pouvant être résumé comme suit.

(i) Tirer une valeur initiale $\theta^{(0)}$ telle que $f(\theta^{(0)} | X) > 0$. Elle est choisie de façon tout à fait quelconque pour autant que cette condition soit respectée. En effet, son influence s'amenuise lorsque la taille de la simulation devient grande. La valeur initiale est d'autant moins importante qu'on élimine parfois les premiers tirages (tirages "burn-in")⁹ de l'analyse.

(ii) Pour $m = 1, 2, \dots, M$

a) Tirer un candidat $\theta^{(*)}$ de la distribution connue à l'itération m étant donné la valeur précédente $\theta^{(m-1)}$. Ce candidat est tiré d'une distribution ("jumping distribution" en référence à Gelman et al. ([8])) que l'on se fixe à l'avance.

b) Calculer le rapport $\alpha = \frac{f(\theta^{(*)}|X)}{f(\theta^{(m-1)}|X)}$ ¹⁰ et appliquer la règle de décision suivante :

-Si $\alpha \geq 1$ on accepte $\theta^{(m)} = \theta^{(*)}$
 -Sinon on compare α à un nombre aléatoire tiré de la loi uniforme sur $[0, 1]$. Si α est supérieur ou égal au nombre tiré de la loi uniforme, alors on prend $\theta^{(m)} = \theta^{(*)}$, sinon on conserve $\theta^{(m)} = \theta^{(m-1)}$.

Sous certaines conditions de régularité dues à Gelman et al. [8], la séquence $\{\theta^{(m)}\}$ ainsi obtenue converge en distribution vers $f(\theta | X)$.

Outre les avantages sus-mentionnés, l'algorithme Metropolis présente la faiblesse d'exiger que la distribution de laquelle sont tirés les paramètres candidats ("jumping distribution") soit symétrique. Cette contrainte est respectée dans notre cas où ladite distribution est normale multivariée. L'algorithme alternatif Metropolis-Hasting relâche cette exigence et emboîte de ce fait l'algorithme Metropolis.

⁹On élimine parfois les 300 à 500 premiers tirages pour une simulation de taille 10 000.

¹⁰Dans le cas d'un algorithme Metropolis-Hasting où la distribution de saut J_m n'est pas symétrique la formule devient $\alpha = \frac{f(\theta^{(*)}|X) / J_m(\theta^{(*)}|\theta^{(m-1)})}{f(\theta^{(m-1)}|X) / J_m(\theta^{(m-1)}|\theta^{(*)})} = \frac{f(\theta^{(*)}|X) J_m(\theta^{(m-1)}|\theta^{(*)})}{f(\theta^{(m-1)}|X) J_m(\theta^{(*)}|\theta^{(m-1)})}$

6.2 Ajustement des données

Les durées entre les transactions observées sur une base journalière exhibent une courbe en U-renversé. C'est à dire qu'elles vont d'abord croissantes, atteignent un maximum en mi-journée et décroissent par la suite. Ceci résulte de plusieurs faits : À l'ouverture de la bourse, les titres s'échangent jusqu'à ce que les nouveaux prix reflètent les informations disponibles depuis la fermeture (fin des activités) de la veille. Ensuite, lorsque la pause de midi approche, il y a le "lunch-effect" qui agit sur les acteurs du marché et ralentit le rythme des activités. Après la pause, les acteurs du marché reprennent leur rythme d'activité normal et les transactions sont donc plus rapprochées. En fin de journée, le fait que les agents veulent fermer leur position accélère davantage le rythme des transactions.

Plusieurs méthodes ont été proposées dans la littérature pour isoler cet effet journalier. Le principe général consiste à construire une fonction déterministe g et de l'appliquer aux données. Les nouvelles durées utilisées pour l'analyse sont alors obtenues par l'équation: $x_i = \frac{\Delta t_i}{g(t_i)}$.

Dans ce travail, nous retenons la fonction g établie par Tsay [20] qui utilise la même base de données IBM que nous. Dans la mesure où les données présentent une forme en U-renversé, il exprime g comme combinaison de fonctions quadratiques négatives de la façon suivante:

$$g(t_i) = \exp(d_i), \quad d(i) = \beta_0 + \sum_{j=1}^7 \beta_j g_j(t_i) \quad (17)$$

où:

$$g_1(t_i) = -\left(\frac{t_i - 43200}{14400}\right)^2, \quad g_2(t_i) = -\left(\frac{t_i - 48300}{9300}\right)^2$$

$$g_3(t_i) = \begin{cases} -\left(\frac{t_i - 38700}{7500}\right)^2 & \text{si } t_i < 43200 \text{ (avant } 12h : 00) \\ 0 & \text{sinon} \end{cases}$$

$$g_4(t_i) = \begin{cases} -\left(\frac{t_i - 48600}{9000}\right)^2 & \text{si } t_i \geq 43200 \text{ (Après } 12h00) \\ 0 & \text{sinon} \end{cases}$$

$g_5(t_i)$ est la variable indicatrice pour les cinq premières minutes suivant l'ouverture des marchés (entre 9h30 et 9h35)

$$\Rightarrow g_5(t_i) = \begin{cases} 1 & \text{si } t_i \leq 34500 \text{ (Avant } 9h35) \\ 0 & \text{sinon} \end{cases}$$

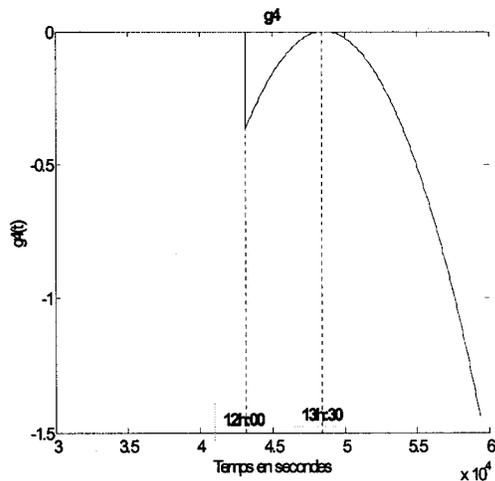
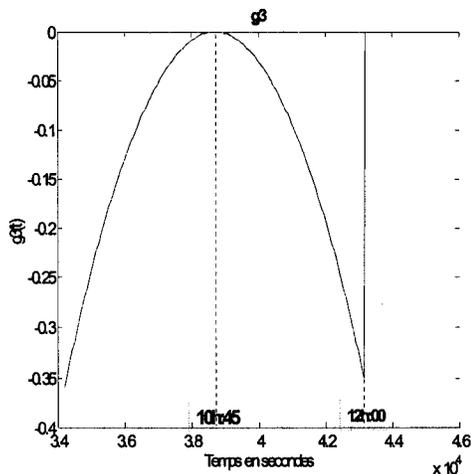
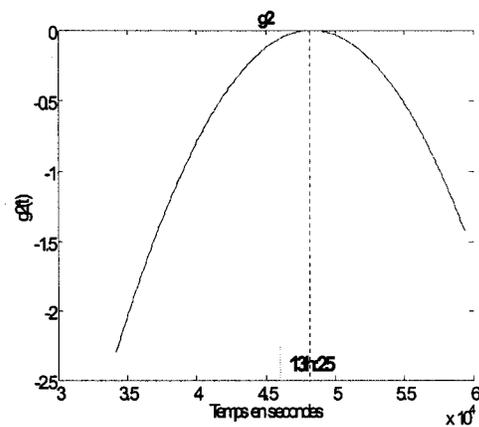
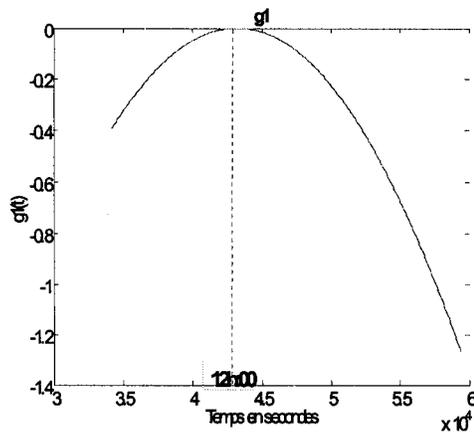
$g_6(t_i)$ est la variable indicatrice pour le deuxième intervalle de cinq minutes d'activité (entre 9h35 et 9h40)

$$\Rightarrow g_6(t_i) = \begin{cases} 1 & \text{si } 34500 \leq t_i \leq 34800 \text{ (Entre 9h35 et 9h40)} \\ 0 & \text{sinon} \end{cases}$$

$g_7(t_i)$ est la variable indicatrice pour les 30 dernières minutes avant la fermeture des marchés (entre 15h30 et 16h00)

$$g_7(t_i) = \begin{cases} 1 & \text{si } t_i \geq 55800 \text{ (Après 15h30)} \\ 0 & \text{sinon} \end{cases}$$

Comme nous pouvons le constater, les fonctions g_j , $j = 1, 2, \dots, 7$ sont définies de façon à caractériser les périodes particulières de la journée.



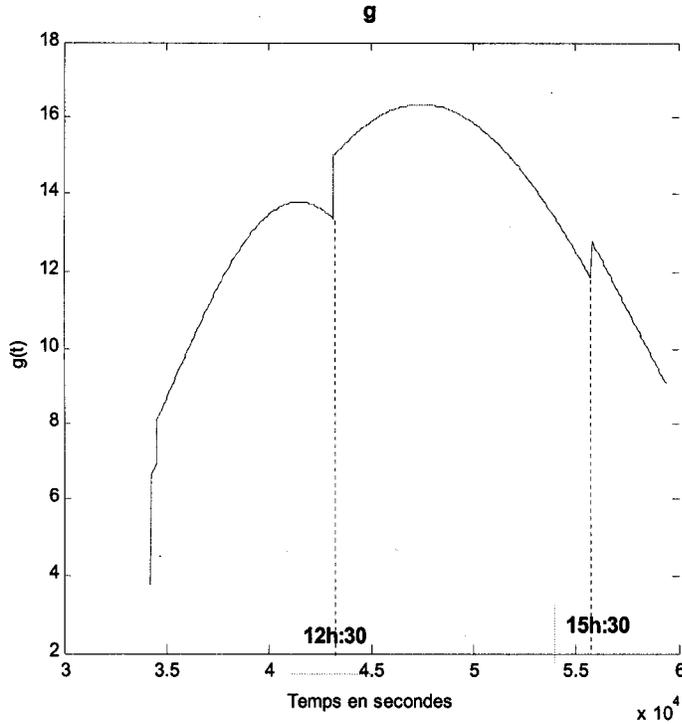
Les coefficients de l'équation (17) sont obtenus par la méthode des moindres carrés ordinaires de la régression linéaire:

$\ln(\Delta t_i) = \beta_0 + \sum_{j=1}^7 \beta_j g_j(t_i) + \epsilon_i$. Le modèle estimé se présente comme suit:

$$\ln(\Delta t_i) = \begin{array}{cccc} 2.813 & +0.171 g_1(t_i) & +0.242 g_2(t_i) & +0.417 g_3(t_i) \\ (0.108) & (0.046) & (0.012) & (0.0849) \\ \\ +0.088 g_4(t_i) & -0.567 g_5(t_i) & -0.151 g_6(t_i) & +0.081 g_7(t_i) \\ (0.059) & (0.044) & (0.042) & (0.034) \end{array}$$

On en déduit donc $g(t_i) = \exp(d_i) = \exp(\ln(\Delta t_i))$ et on procède à l'estimation du modèle proprement dit en utilisant les données débarassées de la composante déterministe.

Ci-dessous la représentation graphique de la fonction estimée g



6.3 Les résultats des estimations

Les résultats obtenus avec l'approche classique du quasi-maximum de vraisemblance¹¹ sont présentés dans le tableau ci-après :

Tableau 1: résultats obtenus par quasi-maximum de vraisemblance

	ω	γ	λ	α
valeur estimée	0.1001	0.0555	0.9048	0.8808
écart-type estimé	0.0046	0.0020	0.0019	0.0112

source: nos estimations

ce qui correspond au modèle estimé suivant:

$$x_i = \eta_i \epsilon_i, \quad \eta_i = 0.1001 + 0.0555x_{i-1} + 0.9048\eta_{i-1}$$

Les résultats des estimations suggèrent que la spécification Weibull serait plus appropriée que la spécification exponentielle pour les innovations. En effet, un test sur le paramètre α révèle qu'il est significativement différent de 1. Nous remarquons également que l'espérance conditionnelle de durée est très persistente au détriment de la durée retardée d'une "période"¹². Ci-après les résultats obtenus avec l'approche bayésienne.

Tableau 2: résultats obtenus par la méthode bayésienne

	ω	γ	λ	α
moyenne à postérieur	0.1263	0.0616	0.8885	0.8787
écart-type à postérieur	0.048	0.0191	0.0310	0.0128

source: nos estimations

Tout comme dans l'approche classique, on note une forte persistance de l'espérance conditionnelle et une évidence en faveur de la distribution Weibull au détriment de la distribution exponentielle pour laquelle α vaudrait 1. Les valeurs moyennes de nos paramètres ne diffèrent pas de façon notable de celles obtenues via l'approche classique du Quasi-ML. Les résultats de l'approche bayésienne suggèrent un effet relativement

¹¹ les estimations ont été faites sous MATLAB version 7.0

¹² période ici fait référence à la transaction précédente et non pas à l'instant précédent comme dans le cas ordinaire où les observations sont régulièrement espacées.

(au Quasi-ML) plus grand de la durée retardée et un effet relativement moins grand de l'espérance conditionnelle retardée déjà très persistente. Cependant, il convient de noter que la moyenne et l'écart-type à posteriori ne sont que quelques éléments de description de la loi à posteriori de nos paramètres. Ils ont été estimés à partir des valeurs simulées des paramètres par application de l'idée de convergence décrite plus haut par la relation (16):

$$\bar{\theta}_k = \frac{1}{10000} \sum_{m=1}^{10000} \left(\theta_k^{(m)} \right) \text{ pour } k = 1, 2, 3, 4.$$

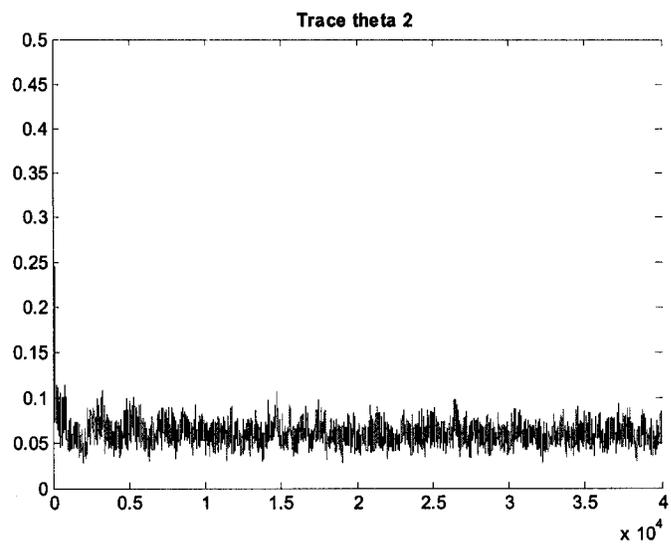
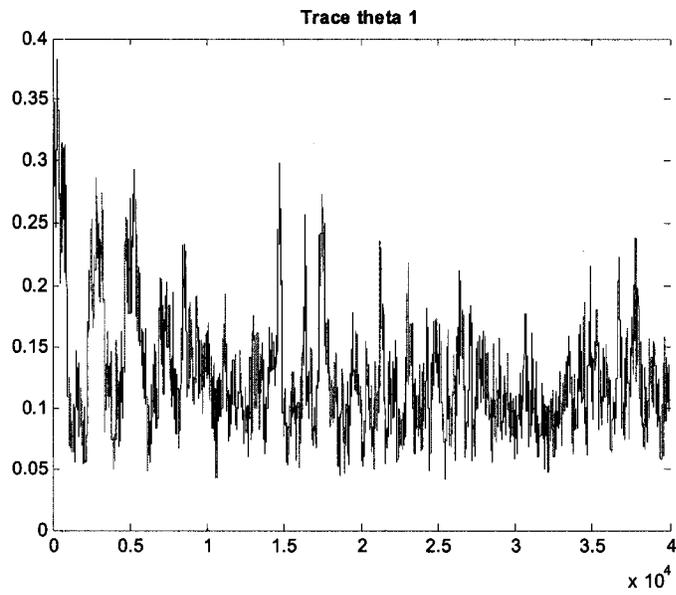
$$\bar{\sigma}_k = \sqrt{\frac{1}{9999} \sum_{m=1}^{10000} \left(\theta_k^{(m)} - \bar{\theta}_k \right)^2} \text{ pour } k = 1, 2, 3, 4$$

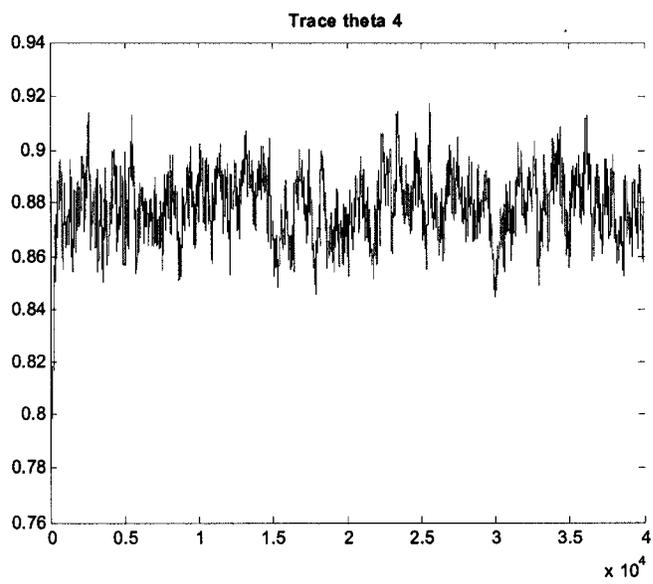
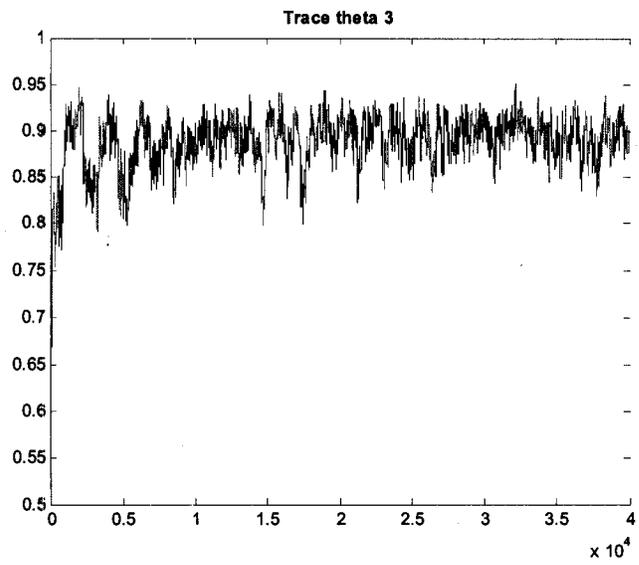
Notons également que les écart-types obtenus sont assez faibles, de l'ordre de 10^{-2}

Les figures ci-après représentent les traces des paramètres à l'exécution de notre programme. Il en ressort principalement que la convergence est assez rapide. Nous en présentons également les densités à posteriori.

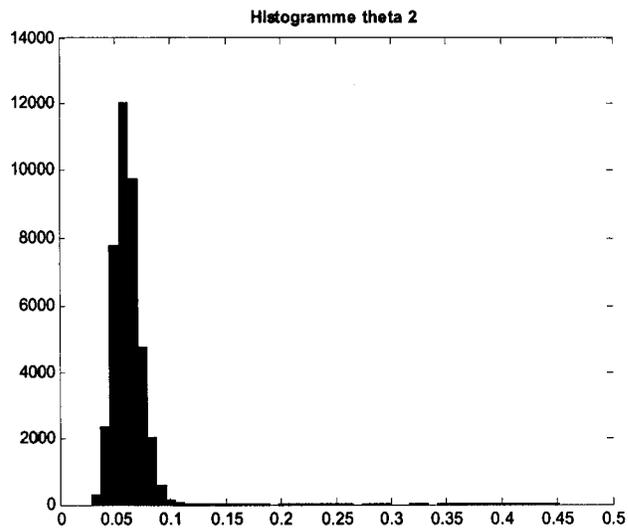
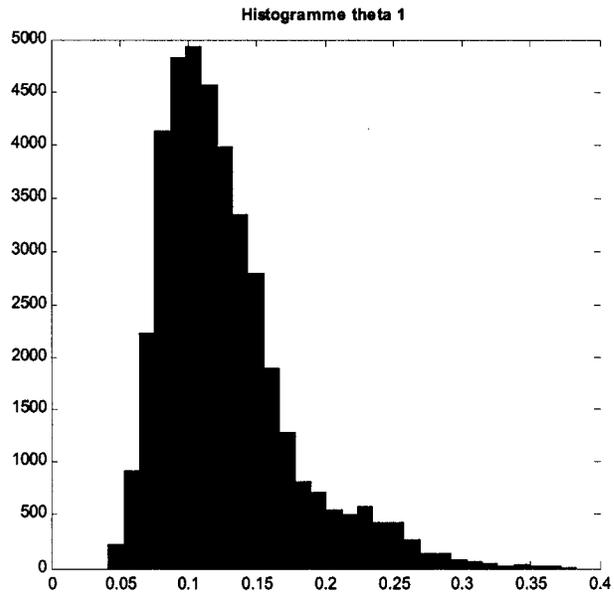
6.3.1 Quelques graphiques des simulations

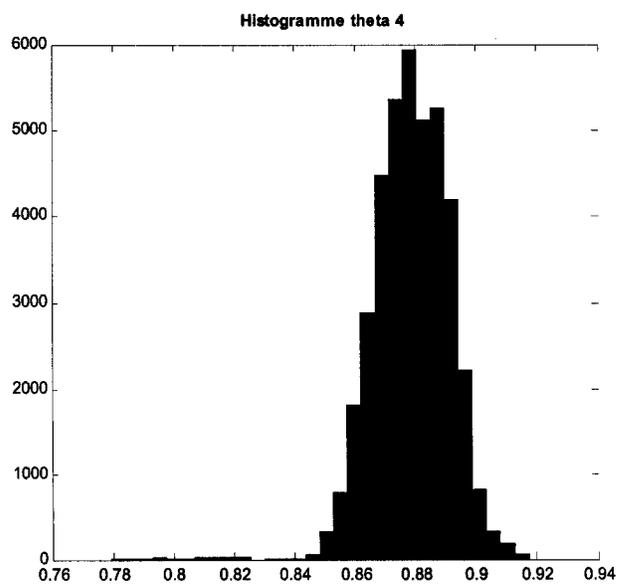
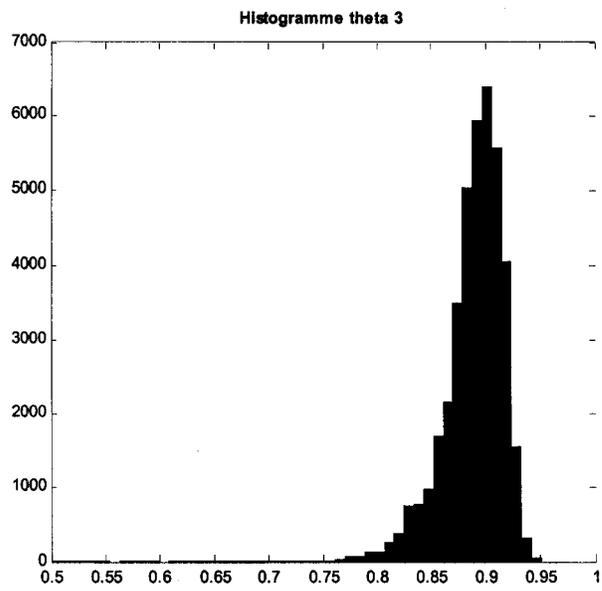
Trace des paramètres $\theta = (\omega, \gamma, \lambda, \alpha)$





Densité à posteriori des paramètres $\theta = (\omega, \gamma, \lambda, \alpha)$





6.4 prévision du nombre de transactions au cours des T prochaines secondes

Le modèle développé précédemment peut-être utilisé à de multiples fins. Nous présentons ici la prévision du nombre de transactions devant avoir lieu dans un intervalle de temps donné. Nous caractérisons cette prévision par un histogramme, exploitant de ce fait un des mérites de l'approche bayésienne adoptée dans ce rapport¹³. Il s'agit pour chaque vecteur de paramètres tiré de notre loi à posteriori d'évaluer le nombre de transactions devant s'effectuer au cours des T prochaines secondes. On peut ainsi tracer un histogramme décrivant la distribution d'une telle variable aléatoire¹⁴. Évidemment nous pouvons également en ressortir quelques statistiques resumées telles que le nombre moyen, l'écart type, le mode,..etc. L'algorithme utilisé est le suivant:

Pour l'initialisation, on se place à l'instant à partir duquel on veut effectuer la prévision. Ceci implique la connaissance de la durée entre les 2 transactions précédentes. On prend ensuite comme valeur estimée de l'espérance de durée conditionnelle la moyenne des durées jusque là observées¹⁵ et on continue comme suit:

- Pour $m = 1, \dots, M$
- (i) Tirer un vecteur de paramètre $\theta = (\omega, \gamma, \lambda, \alpha)$ avec la procédure décrite dans l'algorithme métropolis de la section 5.1
 - (ii) Pour θ tiré en (i), estimer l'espérance de durée conditionnelle par la relation $\eta_i = \omega + \gamma x_{i-1} + \lambda \eta_{i-1}$
 - (iii) Déduire de l'espérance conditionnelle de durée estimée en (ii) le temps écoulé avant la prochaine transaction par la relation $x_i = \eta_i \epsilon_i$ où $\epsilon_i = \frac{\zeta_i}{\beta \Gamma(1 + \frac{1}{\alpha})}$ avec ζ_i tiré d'une loi Weibull de paramètres (α, β) ¹⁶; puis revenir à l'étape (ii).

Tant que le cumul du temps évalué en (iii) est inférieur à T, revenir à l'étape (ii).

¹³Comme on pourra aisément constater dans l'algorithme décrit ci-dessous, une telle prévision n'est pas réalisable avec l'approche classique de l'économétrie.

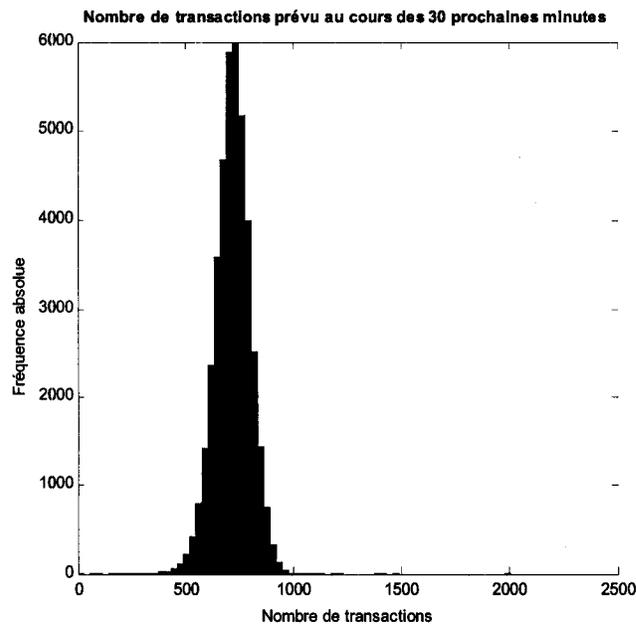
¹⁴C'est effectivement une variable aléatoire vu que le nombre de transactions est évalué pour chaque vecteur de paramètres tiré de façon aléatoire (de la loi à posteriori).

¹⁵Ceci est tout à fait cohérent vu que le terme d'erreur est d'espérance égale à 1 dans notre modèle.

¹⁶notons que le paramètre β n'a en moyenne aucune influence sur les tirages des innovations vu vu le terme de correction $\frac{1}{\beta \Gamma(1 + \frac{1}{\alpha})}$ comme nous l'avons prouvé dans la section réservé à la présentation du modèle théorique.

Une fois qu'on a atteint T , on compte le nombre de transactions enregistrées. Notons que ce nombre est spécifique au tirage d'ordre m du vecteur θ et nous le notons $N(m)$.

L'histogramme de N obtenu pour $T=1800$ secondes (30 minutes) est le suivant:



Le nombre moyen de transactions prévisible à observer en 30 minutes est ainsi de 716 avec un écart-type de 82. Ceci cadre parfaitement avec les données observées. En effet, la durée moyenne entre les transactions est de 2.51 secondes dans l'échantillon, ce qui correspondrait en moyenne à 717 transactions au bout de 30 minutes.

7 Conclusion

Dans cette étude nous nous sommes intéressés à l'analyse bayésienne du modèle ACD: "Autoregressive Conditionnal Duration model" introduit en 1998 par Engle et Russell [7].

Une première étape a consisté à estimer le modèle par la méthode classique du Quasi-maximum de vraisemblance. Dans une deuxième étape, nous avons estimé le modèle en utilisant l'approche bayésienne qui procure l'avantage de prendre en compte l'incertitude qui plane sur les paramètres. En effet, nous partons d'une loi à priori normale multivariée des paramètres pour en dériver la distribution à postériori à la lumière des données. Ces données regorgent les transactions ayant eu cours sur l'action IBM à la bourse de New York entre le 1er novembre 1990 et le 31 janvier 1991. Notre distribution à postériori est simulée en utilisant l'algorithme Metropolis qui visite l'espace des paramètres (de dimension 4 ici) à la recherche de "bons candidats" devant assurer la convergence vers la loi stationnaire.

Les deux approches nous donnent des résultats assez semblables pour ce qui est de la taille des paramètres. Cependant, l'approche bayésienne décrit de façon plus complète la distribution des paramètres. On pourrait par exemple s'intéresser à des moments d'ordre supérieur. Notre approche est ensuite utilisée pour des fins de prévisions, notamment le nombre de transactions dans un intervalle de temps donné. Cette prévision semble bien cadrer avec les données observées vu qu'elle prévoit une moyenne de 710 transactions au bout de 30 minutes, contre 717 dans notre échantillon.

Cette étude a mis en exergue la forte persistance de l'espérance conditionnelle de durées qui conduit la dynamique des transactions, ce qui suggère fortement des pistes d'extension telle que l'utilisation des modèles à mémoire longue.

Par ailleurs, la spécification Weibull s'est révélée mieux appropriée que la spécification exponentielle des innovations pour décrire nos données. Il serait donc intéressant d'envisager d'autres types distributions pour mieux cerner la dynamique des durées qui s'écoulent entre les transactions d'un titre en bourse.

On pourrait également mener des réflexions dans le sens de comparer deux bourses de valeurs en ce qui concerne la dynamique des durées pour un titre transigé dans les deux institutions.

References

- [1] **Bauwens, L., D. Veredas.** (2004). The Stochastic Conditional Duration Model: A Latent Factor Model for the Analysis of Financial Duration, *Journal of Econometrics* 119(2), 381-412.
- [2] **Bauwens, L., Giot, P.**(2000). The Logarithmic ACD Model: an application to the bid-ask quote process of three NYSE stocks. *Annales d'Économie et de la Statistique* 60, 117-149.
- [3] **Campbell, J.Y., Lo, A.W and Mackinlay** (1997), *The Econometrics of Financial Markets*, (Princeton University Press).
- [4] **Dale, J.Poirier, Tobias, J.L.** (2005), *Bayesian Econometrics*, University of California-Irvine & Iowa State University.
- [5] **Easley, D. and O'Hara M.**, (1992). Time and the Process of Security Price Adjustment, *The Journal of Finance*, 47, 905-927.
- [6] **Engle, R.F.** (1982). Autoregressive Conditional Heteroscedasticity with estimates of the variance of United Kingdom Inflation, *Econometrica*, 50, 987-1007.
- [7] **Engle, R.F., Russell, J.R.**, (1998). Autoregressive Conditional Duration: a new approach for irregularly spaced transaction data. *Econometrica* 66, 1127-1162.
- [8] **Gelman, A., Carlin, J.B, Stern, H.S., and Rubin, D.B.** (1995), *Bayesian Data Analysis*, *CRC Press*: London.
- [9] **Grammig, J., Maurer, K.O.** (1999). Non-monotonic Hazard Functions and the Autoregressive Conditional Duration Model, *The econometrics Journal* 3, 16-38.
- [10] **Geweke, J.** (2003), *Contemporary Bayesian Econometrics and Statistics*, Department of economics, University of Iowa.
- [11] **Geweke, J.** (2003), Using Simulation Methods for Bayesian Econometrics Models: Inference, Development and Communication (with discussion), *Econometric Reviews* 18, 1-127.
- [12] **Ghysels, E., C. Gouriéroux, and J. Jasiak.** (2004). Stochastic Volatility Duration Model, *Journal of Econometrics* 119, 413-433.
- [13] **Greene, W.,H.** (2003), *Econometric Analysis*, 5th edition, Prentice Hall.
- [14] **Jaquier, E., Polson, N.G., Rossi, P.** (1994), Bayesian Analysis of Stochastic Volatility Models, *Journal of Business & Economic Statistics*, 12, 371-417.
- [15] **Koop, G.**(2003), *Bayesian Econometrics*, John Wiley & Sons
- [16] **Lancaster T.** (2004), *An Introduction to Modern Bayesian Econometrics*, (Blackwell).
- [17] **Lunde, A.** (1999). *A Generalized Gamma Autoregressive Conditional Duration Model*, Aalborg University.
- [18] **Melino, A., Turnbull, S.M.** (1990), *Pricing Foreign Currency*

options with Stochastic Volatility, *Journal of Econometrics*, 45, 239-265.

- [19] **Simonsen, O.**, (2004), An Empirical Model for Durations in Stocks, Umea University.
- [20] **Tsay, R.S.** (2002), Analysis of Financial Time Series, *Wiley Series in Probabilities and Statistics*.