JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

FRÉDÉRICK BOURGOIN

DAVID PRIEUL

Estimation de diffusion à volatilité stochastique et application au taux d'intérêt à court terme français

Journal de la société statistique de Paris, tome 138, nº 4 (1997), p. 43-66

<http://www.numdam.org/item?id=JSFS_1997__138_4_43_0>

ESTIMATION DE DIFFUSION A VOLATILITÉ STOCHASTIQUE ET APPLICATION AU TAUX D'INTÉRÊT A COURT TERME FRANÇAIS

Frédérick BOURGOIN et David PRIEUL [1]

Résumé

Ce document applique une classe d'estimateurs du khi-deux minimum à l'estimation de systèmes d'équations différentielles stochastiques représentant un modèle structurel. L'observation partielle en temps discret du vecteur d'état du modèle structurel justifie le recours à une méthode d'inférence indirecte simulée. L'espérance de la fonction score d'un modèle auxiliaire estimé par pseudo maximum de vraisemblance est alors utilisée comme critère définissant la fonction objectif de la méthode des moments généralisés. Sous cette approche, le choix du générateur de scores revêt une importance primordiale pour garantir l'efficacité des estimateurs simulés et prévenir le risque d'une mauvaise spécification du modèle structurel. La conclusion principale qui ressort de cette recherche est l'incapacité des modèles traditionnels à un facteur à représenter l'idiosyncrasie de la dynamique du taux d'intérêt à court terme français. L'intégration d'un second facteur non observable relatif à la dynamique aléatoire de volatilité apparaît alors comme une évolution naturelle vers un traitement plus flexible des spécifications en temps continu.

1. Frédérick Bourgoin, Millennium Global Investments, 63 St.James's Street, SW1A 1LY, Londres, E-Mail : fbourgoin@compuserve.com; David Prieul, Lehman Brothers International (Europe), One Broadgate, EC2M 7HA, Londres, E-Mail : dprieul@lehman.com. Ce document de recherche ne reflète pas la position de Millennium Global Investments et Lehman Brothers International (Europe) et n'engage que ses auteurs. Nous remercions Christian Gouriéroux pour ses commentaires sur des versions précédentes de l'article.

1. Introduction

Depuis l'introduction par MERTON [13] du concept de processus aléatoire en temps continu pour la modélisation de la dynamique du niveau de taux d'intérêt nominal sans risque instantané, de nombreux travaux se sont attachés à généraliser le modèle initial. On peut à ce titre citer le modèle de VASICEK [15], qui réduit de façon substantielle la probabilité d'obtenir des taux d'intérêt nominaux négatifs (le taux d'intérêt instantané suit alors un processus markovien d'Ornstein-Uhlenbeck à variance constante et tend à régresser vers une valeur moyenne de longue période), et le processus racine carrée de COX, INGERSOLL, et ROSS [5] qui satisfait à la contrainte de non négativité du taux d'intérêt nominal tout en représentant la volatilité du taux d'intérêt instantané comme une fonction croissante du niveau de celui-ci (introduisant par là même une forme d'hétéroscédasticité).

L'ensemble des modélisations en temps continu du taux instantané peuvent se généraliser sous la forme d'un générateur aléatoire de données ou équation différentielle stochastique dérivé par CHAN, KAROLYI, LONGSTAFF et SANDERS [4] (CKLS)

$$\mathrm{d}r_t = (\alpha + \beta r_t)\mathrm{d}t + \sigma r_t^\gamma \mathrm{d}W_t \ ; \ 0 \leq t < T \tag{1.1}$$

où r_t représente le taux d'intérêt instantané ou vecteur d'état, $W = \{W_t, t \geq 0\}$ est un processus de Wiener standard, et α, β, σ et γ sont des paramètres. L'évolution moyenne du taux instantané est représentée par le terme de tendance $(\alpha + \beta r_t)$ tandis que $\sigma^2 r_t^{2\gamma}$ désigne la variance non anticipée du changement instantané de r_t. Le paramètre de volatilité σ peut, dans un premier temps, être interprété comme un simple facteur d'échelle, tandis que γ représente le lien existant entre volatilité et niveau de taux d'intérêt. Différentes contraintes posées sur la valeur des paramètres permettent d'obtenir les modèles de Merton ($\beta = \gamma = 0$), de Vasicek ($\gamma = 0$), de BRENNAN et SCHWARTZ [1] ($\gamma = 1$) et de Cox, Ingersoll, et Ross ($\gamma = 0.5$). L'objet de l'étude est alors d'estimer l'ensemble des paramètres de (1.1). En d'autres termes, si l'on dénomme par $\{r_t\}_{t=0}^T$ une réalisation de (1.1) et $\{\tilde{r}_t\}_{t=1}^n$ une suite d'observations en temps discret du processus en temps continu, avec $\tilde{r}_t = r_{tk}$, $t = (1, \ldots, n)$, $nk \leq T$ quel que soit $k > 0$, le problème se résume à l'estimation du vecteur de paramètres $\rho = (\alpha, \beta, \gamma, \sigma)$ du modèle structurel. La loi du processus observé est généralement explicitée en fonction de la probabilité de transition ou densité conditionnelle $p(r_{t+1} \,|\, r_t, \rho)$. Une approche du type maximum de vraisemblance est alors appliquée pour estimer ρ. Cependant, la forme fonctionnelle de cette densité n'est connue que pour des cas particuliers [2], nécessitant le recours à des méthodes alternatives d'estimation

2. Ainsi, on montre que $p(r_{t+1} \,|\, r_t, \rho)$ suit une loi du khi-deux non centré si l'évolution moyenne du taux instantané est linéaire et la fonction de diffusion est un processus racine carrée. La loi marginale du processus est alors une loi gamma. Dans le cas d'un processus d'Ornstein-Uhlenbeck, les observations peuvent être représentées par une densité de probabilité gaussienne. L'estimation des paramètres d'intérêt s'effectue alors par maximum de vraisemblance dynamique.

dans le cas où une discrétisation du modèle structurel aboutit à un modèle d'espace d'état non linéaire.

L'examen des diverses études empiriques relatives à l'estimation de spécifications du type (1.1) en temps discret révèle qu'aucun consensus n'apparaît véritablement quant à la modélisation de la dynamique de niveau et de volatilité du taux d'intérêt instantané. Comme le soulignent CKLS, l'absence de contrainte sur γ conduit à retenir une valeur estimée de ce paramètre supérieure à l'unité. Les résultats présentés par BROZE, SCAILLET, et ZAKOIAN [3] prouvent également que la stabilité dynamique d'une approximation en temps discret de (1.1) dépend de la valeur de γ estimée. Enfin, BRENNER, HARJES, et KRONER [2] (BHK) critiquent le fait que l'approximation en temps discret utilisée par CKLS ne permet pas à des chocs non anticipés de taux d'agir sur la dynamique de volatilité ; cela restreint cette dernière à n'être alors qu'une simple fonction du niveau du taux d'intérêt. Aussi, BHK montrent que l'introduction d'une spécification autorégressive à hétéroscédasticité conditionnelle (ARCH) met en évidence un important degré de persistance dans la variance conditionnelle du processus de taux d'intérêt et se traduit par des valeurs estimées de γ non significativement différentes de 0.5. Leur préférence se porte alors sur une spécification paramétrique mixte fournissant une première approximation de la distribution conditionnelle du changement de r_t où niveau et innovation de taux d'intérêt modélisent la dynamique de volatilité.

L'ensemble de ces résultats met également en évidence un certain nombre de problèmes liés à l'estimation d'une diffusion de taux d'intérêt. La présence d'un biais de discrétisation amène à douter de la pertinence du lien entre paramètres discrets et continus. La stabilité des paramètres et de la dynamique des spécifications estimées est souvent ignorée, ce qui entraîne des problèmes lors de l'évaluation d'actifs dont le prix dépend de sentiers d'évolution simulés. Enfin, le traitement de la question de la cohérence entre la densité conditionnelle estimée représentative de la série chronologique à temps discret de taux d'intérêt et une suite de réalisations discrètes de (1.1) est généralement absent.

Cet article présente une application de la méthode d'inférence indirecte des moments efficaces (MME) de GALLANT et TAUCHEN [10] à l'estimation de systèmes d'équations différentielles stochastiques. L'estimateur obtenu peut être assimilé à un estimateur des moments généralisés dont les propriétés asymptotiques sont similaires à celles d'un estimateur du maximum de vraisemblance, lorsqu'un nombre suffisant de conditions de moments et d'observations sont disponibles. La méthode MME propose une approche systématique pour le choix des moments définissant la fonction objectif de la méthode des moments généralisés (MMG). L'espérance de la fonction score d'un modèle auxiliaire (appelé générateur de scores) estimé par pseudo maximum de vraisemblance est alors utilisée comme critère définissant la fonction objectif de la méthode MMG. Enfin, la qualité de l'ajustement peut être directement testée. La statistique du khi-deux dérivée de l'estimation par MMG peut s'interpréter comme un test global de bonne spécification du modèle structurel. L'analyse des composantes de Student du critère du khi-deux permet également d'obtenir une série de diagnostics sur les propriétés du modèle structurel qui ne conviennent pas au regard du modèle auxiliaire estimé.

La section 2 présente les principes généraux de la méthode MME. Les propriétés statistiques des données retenues pour l'étude empirique sont étudiées en section 3. Une attention toute particulière est portée sur l'estimation de la densité de probabilité conditionnelle représentative de l'échantillon observé, *i.e.*, le modèle auxiliaire. La section 4 applique la méthode MME à l'estimation de diffusions à volatilité stochastique relatives au taux d'intérêt à court terme français. La dernière section conclut.

2. Méthode des moments efficaces et diffusions

2.1 Motivation

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbf{P})$ un espace probabilisé muni d'une filtration. Les méthodes d'évaluation d'actifs contingents représentent habituellement les variables d'état sous la forme de processus de diffusion en temps continu, satisfaisants au système d'équations différentielles stochastiques suivant

$$\mathrm{d}X_t = \mu(t, X_t, \rho)\mathrm{d}t + \sigma(t, X_t, \rho)\mathrm{d}W_t \ ; \ 0 \leqslant t < T \tag{2.1}$$

où $\rho \in \mathbb{R}^{p_\rho}$ est un vecteur de paramètres de dimension p_ρ, X_t est un vecteur d'état de dimension s, $W = \{W_t,\ t \geqslant 0\}$ représente un mouvement brownien w-dimensionnel à valeur dans $\mathbb{R}^w$ adapté à $\mathcal{F}_t$, où les $(W_t)_{t\geqslant 0}$ sont des $\mathcal{F}_t$-mouvements browniens standards indépendants- et où $\mu(\cdot,\cdot,\rho) : [0,T]\times\Omega \to \mathbb{R}^s$ et $\sigma(\cdot,\cdot,\rho) : [0,T]\times\Omega \to \mathbb{R}^{s\times w}$ représentent respectivement l'évolution moyenne et la diffusion du processus. Le système (2.1) peut être assimilé à un générateur aléatoire de données représentatif d'un modèle structurel, avec $\{X_t\}_{t=0}^T$ représentant une solution du processus d'Itô multidimensionnel suivant

$$X_t = X_0 + \int_0^t \mu(u, X_u, \rho)\,\mathrm{d}u + \sum_{i=1}^{w} \int_0^t \sigma_i(u, X_u, \rho)\,\mathrm{d}W_{iu} \tag{2.2}$$

$$\mu_t \text{ et les } \sigma_{it} \text{ sont adaptés à } \mathcal{F}_t$$

$$\int_0^T |\,\mu(u, X_u, \rho)\,|\,\mathrm{d}u < +\infty \ \mathbf{P}\ p.s.$$

$$\int_0^T \left[\sigma_i(u, X_u, \rho)\right]^2 \mathrm{d}W_{iu} < \ +\infty \ \mathbf{P}\ p.s.$$

où X_0 est la condition initiale au temps $t = 0$ (un échantillon de la distribution stationnaire), et $\int_0^t \sigma_i(u, X_u, \rho)\,\mathrm{d}W_{iu}$ représente l'intégrale stochastique d'Itô.

Etant donnés $\mu(\cdot,\cdot,\rho)$ et $\sigma(\cdot,\cdot,\rho)$, la densité de transition $p(Y'_\tau|Y_t)$ d'une valeur Y à la date t à une valeur Y' à la date $\tau, \tau > t$, satisfait alors les équations forward et backward de Kolmogorov

$$\frac{\partial p(Y'_\tau|Y_t)}{\partial \tau} + \frac{\partial}{\partial Y'}\left[\mu(Y')p(Y'_\tau|Y_t)\right] - \frac{1}{2}\frac{\partial^2}{\partial (Y')^2}\left[\sigma^2(Y')p(Y'_\tau|Y_t)\right] = 0$$

pour (t, Y) fixés et

$$\frac{\partial p(Y'_\tau\,|\,Y_t)}{\partial t} + \mu(Y)\frac{\partial}{\partial Y}\left[p(Y'_\tau\,|\,Y_t)\right] + \frac{1}{2}\sigma^2(Y)\frac{\partial^2}{\partial Y^2}\left[p(Y'_\tau\,|\,Y_t)\right] = 0$$

pour (τ, Y') fixés. Pour des spécifications paramétriques particulières de $\mu(\cdot,\cdot,\rho)$ et $\sigma(\cdot,\cdot,\rho)$, la résolution de ces équations permet de représenter la densité conditionnelle comme une fonction des paramètres d'intérêt, et d'utiliser une approche par maximum de vraisemblance pour estimer les paramètres du modèle structurel. Néanmoins, ces équations ne peuvent généralement se résoudre que numériquement, rendant l'estimation par maximum de vraisemblance extrêmement difficile, voire impossible dans le cas d'un vecteur d'état partiellement observable.

2.2 Méthode des moments efficaces

Un certain nombre de méthodes alternatives ont alors été proposées pour l'estimation des paramètres du système d'équations (2.1), en particulier l'estimation par MMG avec une suite de conditions de moments d'ordre faible approximés en temps discret. Ainsi, CKLS estiment leur modèle de taux d'intérêt à l'aide d'une discrétisation d'Euler-Maruyama de (1.1)

$$r_{t+\Delta t} - r_t = \alpha + \beta r_t + \varepsilon_{t+\Delta t} \tag{2.3}$$

avec $E_t(\varepsilon_{t+\Delta t}) = 0$, $E_t(\varepsilon^2_{t+\Delta t}) = \sigma^2 r_t^{2\gamma}$, $t = (1, \ldots, n)$, et $\rho = (\alpha, \beta, \sigma, \gamma)$. Un estimateur de ρ déterminé tel que

$$\hat{\rho} = \operatorname{argmin}_\rho \left[\left[\frac{1}{n}\sum_{t=1}^{n} f_t(\rho)\right]' \mathbf{W}_n \left[\frac{1}{n}\sum_{t=1}^{n} f_t(\rho)\right]\right]$$

où la matrice de pondération $\mathbf{W}_n$, symétrique définie positive, est déterminée de façon optimale, et où $f_t(\rho)$ représente un vecteur de moments conditionnels d'ordre faible. Les inconvénients de cette méthode sont nombreux : l'approximation en temps discret n'est correcte que si $\Delta t \to 0$, aucun guide n'est fourni quant à la nature et le nombre de conditions de moments entrant dans la définition de $f_t(\rho)$, et les facteurs implicites à la spécification en temps continu doivent être observables.

L'approche poursuivie ici se présente comme une alternative aux représentations précédentes en proposant une méthode d'estimation générale et efficace d'une diffusion de taux d'intérêt. Si l'on représente par X_t le vecteur d'état partiellement observable d'un processus générateur de Markov strictement stationnaire et ergodique, de vecteur de paramètres $\rho \in \mathbb{R}^{p_\rho}$, on définit alors une fonction observable du vecteur d'état, de distribution absolument continue, sous la forme $z_t = g(t, X_t, \rho)$, de densité homogène $p(z_{t-L}, \ldots, z_t \mid \rho)$, avec

$$p(z_t \mid z_{t-J}, \ldots, z_{t-1}, \rho) = \frac{p(z_{t-J}, \ldots, z_t \mid \rho)}{p(z_{t-J}, \ldots, z_{t-1} \mid \rho)}, \quad J = 1, 2, \ldots$$

représentant les densités de transition. De façon générale, on suppose que le processus observé est markovien, tel que

$$p(z_t \mid z_{t-J}, \ldots, z_{t-1}, \rho) \approx p(z_t \mid z_{t-L}, \ldots, z_{t-1}, \rho), \quad J \geq L$$

pour L grand. Sous cette approche, les densités jointes et de transition ne possèdent pas de forme analytique immédiate. La méthode MME repose premièrement sur la définition d'un modèle auxiliaire, *i.e.*, une densité conditionnelle pour la suite de variables aléatoires z_t, de forme $f(z_t \mid z_{t-L}, \ldots, z_{t-1}, \theta), \theta \in \mathbb{R}^{p_\theta}$. Le modèle auxiliaire est sélectionné de façon à représenter au mieux le processus générateur des observations $\{\tilde{z}_t\}_{t=1}^n$. En d'autres termes, le modèle auxiliaire $f(z_t|z_{t-L} \ldots, z_{t-1}, \theta)$ est considéré comme une bonne approximation de la densité de transition $p(z_t|z_{t-L}, \ldots, z_{t-1}, \rho)$. L'ajustement du modèle auxiliaire aux données $\tilde{z}_t$ est effectué par pseudo maximum de vraisemblance afin d'obtenir un estimateur de θ

$$\tilde{\theta} = \operatorname{argmax}_\theta \left[\frac{1}{n}\sum_{t=1}^{n} \log\left[f(\tilde{z}_t \mid \tilde{z}_{t-L}, \ldots, \tilde{z}_{t-1}, \theta)\right]\right] \tag{2.4}$$

et un estimateur de la matrice d'information $\mathcal{I}_n$

$$\mathcal{I}_n = \frac{1}{n}\sum_{t=1}^{n} \left[\frac{\partial}{\partial\theta}\log\left[f(\tilde{z}_t|\tilde{z}_{t-L}, \ldots, \tilde{z}_{t-1}, \theta)\right]\right]\left[\frac{\partial}{\partial\theta}\log\left[f(\tilde{z}_t|\tilde{z}_{t-L}, \ldots, \tilde{z}_{t-1}, \theta)\right]\right]'$$

La condition du premier ordre pour l'estimation de θ est donnée par

$$\frac{1}{n}\sum_{t=1}^{n} \frac{\partial}{\partial\theta}\log\left[f(\tilde{z}_t|\tilde{z}_{t-L}, \ldots, \tilde{z}_{t-1}, \tilde{\theta})\right] = 0$$

Les conditions sur les moments de la méthode MME sont, dans un second temps, définies comme l'espérance, sous l'hypothèse d'un modèle structurel, de la fonction score du modèle auxiliaire

$$m(\rho, \theta) = \int \cdots \int \frac{\partial}{\partial\theta}\log\left[f(z_t|x_t, \theta)\right] p(x_t, z_t \mid \rho) \prod_{k=0}^{L} \mathrm{d}z_{t-k} \tag{2.5}$$

avec $x_t = (z_{t-L}, \ldots, z_{t-1})$. Les conditions de moments dépendent alors des paramètres du modèle auxiliaire et de ceux du modèle structurel. La présence de l'intégrale dans l'équation (2.5) rend l'espérance non calculable analytiquement, et la méthode d'intégration Monte Carlo est utilisée à la place. Le calcul de ces moments s'effectue alors en générant une longue simulation $\{\hat{z}_t\}_{t=1}^N$ du modèle structurel et en prenant la moyenne des scores

$$m(\rho,\theta) \approx m_N(\rho,\theta)$$

$$m_N(\rho,\theta) = \frac{1}{N}\sum_{t=1}^{N} \frac{\partial}{\partial\theta}\log\left[f(\hat{z}_t|\hat{z}_{t-L},\ldots,\hat{z}_{t-1},\theta)\right]$$

où $\approx$ représente l'équivalence asymptotique, $N \geq 50\,000$ de façon à minimiser l'erreur Monte Carlo, et où $m_N(\rho,\theta)$ est utilisée à la place de $m(\rho,\theta)$, le modèle structurel étant supposé stationnaire et ergodique. Chaque simulation est dérivée pour une valeur particulière du vecteur de paramètres ρ. On représente alors une simulation basée sur ρ et sur l'histoire du processus simulé $\left(\hat{z}_{t-L}(\rho),\ldots,\hat{z}_{t-1}(\rho)\right)$ sous la forme $\{\hat{z}_t(\rho)\}_{t=1}^N$. Les conditions sur les moments peuvent alors se récrire en tenant compte explicitement de la nature de la dépendance en ρ

$$m_N(\rho,\theta) = \frac{1}{N}\sum_{t=1}^{N} \frac{\partial}{\partial\theta}\log\left[f(\hat{z}_t(\rho)|\hat{z}_{t-L}(\rho),\ldots,\hat{z}_{t-1}(\rho),\theta)\right]$$

On reproduit alors la condition du premier ordre utilisée lors de l'estimation du modèle auxiliaire sous la forme d'une condition de moment dans la simulation en évaluant ρ de telle sorte que

$$m_N(\rho,\tilde{\theta}) = \frac{1}{N}\sum_{t=1}^{N} \frac{\partial}{\partial\theta}\log\left[f(\hat{z}_t(\rho)|\hat{z}_{t-L}(\rho),\ldots,\hat{z}_{t-1}(\rho),\tilde{\theta})\right] \approx 0 \qquad (2.6)$$

où la vraie valeur de θ est remplacée par son estimateur du pseudo maximum de vraisemblance $\tilde{\theta}$. Le critère de la méthode MMG basé sur la métrique $\mathcal{I}_n^{-1}$ minimise alors la distance de (2.6). L'estimateur du khi-deux minimum de ρ est représenté par la statistique

$$\hat{\rho} = \operatorname{argmin}_\rho\left[m_N(\rho,\tilde{\theta})'(\tilde{\mathcal{I}}_n^{-1})m_N(\rho,\tilde{\theta})\right] \qquad (2.7)$$

avec $p_\rho < p_\theta$. Le calcul de $\hat{\rho}$ ne nécessite pas d'émettre de restrictions relatives à l'espace des paramètres pour garantir la stationnarité du processus, celle-ci étant automatiquement assurée lors du calcul des conditions de moments. Ainsi, lorsque les paramètres du modèle structurel estimés à l'aide du modèle auxiliaire ne se situent pas dans une région de stationnarité, les simulations dérivées du générateur aléatoire de données sont explosives et donnent lieu à des valeurs élevées de la fonction objectif à minimiser (TAUCHEN [14]).

2.3 Propriétés et inférence

La théorie MMG s'appliquant, l'estimateur MME est optimal si la matrice de pondération est un estimateur convergent de l'inverse de la matrice de covariance asymptotique des conditions de moments. Si le générateur de scores ajuste correctement la distribution conditionnelle, les scores présentent un comportement proche d'une différence de martingale, et la matrice de covariance de $\sqrt{n}m(\rho,\tilde{\theta})$ peut être estimée à partir de $\tilde{\mathcal{I}}_n$. Si l'on considère que le modèle auxiliaire s'approche de façon grossière du processus sous-jacent aux données observées, un estimateur pondéré de la matrice de covariance asymptotique peut être utilisé pour $\tilde{\mathcal{I}}_n$. A ce stade, il apparait néanmoins préférable de retravailler sur la spécification du modèle auxiliaire, celui-ci se prêtant plus facilement à ce type d'ajustement. La qualité de l'ajustement de la densité conditionnelle par le modèle auxiliaire détermine le degré d'efficacité de l'estimateur. Plus le générateur de scores est une bonne approximation du processus sous-jacent aux données, plus la matrice de covariance asymptotique est proche de celle du maximum de vraisemblance.

Sous des conditions de régularité suffisantes, avec θ^o une solution des équations $m(\rho^o,\theta)=0$, et ρ^o représentant la vraie valeur de ρ dans le système (2.1) et dans la densité $p(x_t, z_t|\rho)$, GALLANT et TAUCHEN [10] montrent que $\hat{\rho}$ est un estimateur sans biais, convergent *p.s.* vers ρ^o, avec $\sqrt{n}(\hat{\rho}-\rho^o)$ convergent en distribution vers $\mathcal{N}(0,\{(M^o)'(\mathcal{I}^o)^{-1}(M^o)\}^{-1})$, où $M^o = M(\rho^o,\theta^o), M(\rho,\theta) = (\partial/\partial\rho')m(\rho,\theta)$, et

$$\mathcal{I}^o = \int\cdots\int \mathbf{A}\mathbf{A}' p(x_t, z_t \mid \rho^o)\prod_{k=0}^{L} \mathrm{d}z_{t-k}$$

avec

$$\mathbf{A} = (\partial/\partial\theta)\log\left[f(z_t \mid x_t, \theta^o)\right]$$

Un estimateur de M^o est représenté par $\widehat{M} = M(\hat{\rho},\tilde{\theta})$, avec $\lim_{n\to\infty}\widehat{M} = M^o$, et $\lim_{n\to\infty}\tilde{\mathcal{I}}_n = \mathcal{I}^o$. Sous l'hypothèse nulle de spécification correcte du modèle structurel $p(x_t, z_t \mid \rho)$, la valeur minimisée de la fonction critère MME, définie par

$$n m_N(\rho,\tilde{\theta})'(\tilde{\mathcal{I}}_n^{-1}) m_N(\rho,\tilde{\theta})$$

est distribuée comme un khi-deux à $(p_\theta - p_\rho)$ degrés de liberté. Dans le cas où l'hypothèse nulle de spécification globale est rejetée, une étude des éléments individuels de Student du vecteur de scores $m_N(\hat{\rho},\tilde{\theta})$ permet d'obtenir un diagnostic sur les caractéristiques du modèle structurel qui ne conviennent pas au regard du modèle auxiliaire estimé. La distribution limite de $\sqrt{n}m(\hat{\rho},\tilde{\theta})$ étant $\mathcal{N}\big(0,\mathcal{I}^o - (M^o)\{(M^o)'(\mathcal{I}^o)^{-1}(M^o)\}^{-1}(M^o)'\big)$, il en découle alors deux types de ratios de Student. Le premier permet d'éviter le calcul des dérivées des scores et permet d'avoir un rapide aperçu de la qualité de l'estimation,

$$\widehat{Q} = \left[\mathrm{diag}(\tilde{\mathcal{I}}_n)\right]^{\frac{1}{2}} \sqrt{n}\, m(\hat{\rho},\tilde{\theta}) \tag{2.8}$$

De larges valeurs de $\widehat{Q}$ reflètent les parties du modèle structurel qui ne peuvent être assimilées aux données. Ces ratios de Student sont néanmoins biaisés vers le bas. De faibles valeurs permettent donc difficilement de conclure quant à la significativité statistique des scores. Leur utilité à des fins d'inférence est donc restreinte. Le second type de ratio corrige le biais en calculant numériquement le jacobien $\widehat{M}$ et en remplaçant $\mathrm{diag}(\tilde{\mathcal{I}}_n)$ par $\mathrm{diag}\big(\tilde{\mathcal{I}}_n - \widehat{M}\{\widehat{M}'(\tilde{\mathcal{I}}_n)^{-1}\widehat{M}\}^{-1}\widehat{M}'\big)$ dans (2.8).

2.4 Simulation

L'application de la méthode MME à l'estimation de (2.1) est cependant délicate à mener à deux points de vue. Le premier reflète la nécessaire stabilité dynamique du modèle auxiliaire, i.e., le générateur markovien de scores. L'utilisation d'un modèle auxiliaire non stationnaire conduit en effet à d'importants problèmes numériques lors de l'étape d'estimation par simulation. Un grand soin doit donc être apporté lors de l'ajustement du modèle auxiliaire aux données, avec en particulier la vérification par simulation de la stabilité du générateur de scores.

Le second point est lié à la nécessaire représentation de (2.1) sous la forme d'un générateur aléatoire de données afin de simuler une série de réalisations discrètes d'un processus en temps continu. Afin de minimiser tout biais de discrétisation dans la simulation, nous appliquons une solution approximée faible de la solution exacte de l'équation différentielle stochastique. Cette approximation valide pour des systèmes autonomes augmente le schéma de discrétisation élémentaire d'Euler par un terme d'Itô-Taylor et converge plus rapidement lorsque l'intervalle de discrétisation tend vers zéro (KLOEDEN et PLATEN [12], Section 15.1).

Une simulation $\{\hat{z}_t\}_{t=1}^{N}$ est déterminée pour un échantillon de données hebdomadaires en divisant chaque semaine en S intervalles, avec S =14 (une demi-journée) pour les applications empiriques[3]. Le pas de discrétisation $\Delta = 1/S$ permet de générer une suite $\widehat{X}_{i\Delta}$ pour $i = \big(1,\ldots,SN)\big)$ et d'obtenir $\hat{z}_{t-L} = g(\widehat{X}_{\Delta(S,t)}) = g(\widehat{X}_t)$ pour $t = (1,\ldots,N)$.

2.5 Diffusion de taux d'intérêt

Dans le contexte de l'estimation d'une diffusion de taux d'intérêt, la spécification (2.1) peut se récrire sous la forme d'un modèle à deux facteurs, où $X_{1,t}$ représente le taux d'intérêt nominal sans risque instantané, et $X_{2,t}$ caractérise un second facteur non observable relatif à la dynamique aléatoire de volatilité

3. Les nombres pseudo aléatoires sont générés à l'aide de la méthode des congruences linéaires de Park et Miller avec réarrangement de Bays-Durham (voir FLANNERY, PRESS, TEUKOLSKY et VETTETLING [7]). La simulation de variables aléatoires gaussiennes centrées et réduites reprend la méthode du rapport de deux variables aléatoires uniformes de KINDERMAN et MONAHAN [11].

$$\mathrm{d}X_{1,t} = (t_{11} + t_{12}X_{1,t})\mathrm{d}t + (v_{11} + v_{12}X_{1,t})^{\gamma} \exp(X_{2,t})\mathrm{d}W_{1,t} \tag{2.9}$$

$$\mathrm{d}X_{2,t} = (t_{21} + t_{22}X_{2,t})\mathrm{d}t + (v_{21} + v_{22}X_{2,t})\mathrm{d}W_{2,t} \tag{2.10}$$

où $W_{1,t}$ et $W_{2,t}$ représentent deux mouvements browniens standards indépendants. Le modèle représenté par les équations (2.9) et (2.10) est identifié sous le nom de diffusion à volatilité stochastique (DVS). Le tableau 1 reprend l'ensemble des modèles qui sont généralisés par la représentation DVS. Dans le cas d'un modèle à un facteur représenté sous la forme d'un mouvement brownien géométrique ($\gamma = 1$) avec tendance et fonction de diffusion intégrant une ordonnée à l'origine non nulle (modèle MBG-plus), la fonction de diffusion se dissipe en $X_1^B = -(v_{11}/v_{12})$ et détermine une barrière réfléchissante à droite ou à gauche selon que la fonction de tendance évaluée en X_1^B soit respectivement positive ou négative. De plus, $-(t_{11}/t_{12})$ définit un état stationnaire pour le taux d'intérêt, *i.e.*, la moyenne non conditionnelle. L'estimation du paramètre γ génère une dynamique d'hétéroscédasticité en fonction du niveau du taux d'intérêt.

3. Analyse empirique

3.1 Nature des données

L'échantillon utilisé pour l'estimation d'une diffusion de taux d'intérêt consiste en 1 178 observations hebdomadaires du taux d'intérêt du marché monétaire français à trois mois composé continûment sur la période allant du 18 janvier 1974 au 9 août 1996, $\{\tilde{z}_t\}_{t=1}^{1178}$. L'analyse des données souligne un important degré de persistance présent dans la série de taux d'intérêt, doublé de valeurs extrêmes associées à des chocs exogènes. La série de taux en différences premières présente 10 observations au-delà de 10 écarts types qui sont associées à des chocs politiques ou de change. L'existence de tels chocs et la présence d'une forte persistance dans la moyenne conditionnelle peuvent conduire à des problèmes de stabilité numérique lors d'estimations faisant appel à des procédures récursives et incite donc à une grande prudence lors du choix de la forme fonctionnelle et de l'estimation des paramètres du modèle auxiliaire.

La moyenne empirique sur l'échantillon est de 9,25 %, avec un écart-type de 2,55 %. Les données en différences premières ont une moyenne de -0,0085 % et un écart-type de 0,295 %. Le maximum enregistré sur la période est de 17,92 % le 12 juillet 1981, le minimum étant de 3,62 % le 19 juillet 1996. L'augmentation la plus forte sur une semaine est de 2,65 % le 22 mai 1981, et la plus forte baisse est de -1,70 % la semaine du 10 septembre 1981. La figure 1 présente l'évolution du taux du marché monétaire au cours de la période étudiée.

Le tableau 2 résume la structure statistique des données en niveau $\tilde{z}_t$ et en différences premières $(\tilde{z}_t - \tilde{z}_{t-1})$. L'examen des fonctions d'autocorrélations empiriques confirme une extrême persistance dans la série de taux d'intérêt, et valide la présence d'une quasi-tendance aléatoire dans $\tilde{z}_t$. La série en différences premières présente les caractéristiques d'un bruit blanc faible non gaussien. Les statistiques portmanteau de Ljung-Box (LB) ne sont pas significatives aux seuils statistiques usuels et soulignent l'absence de toute structure de dépendance linéaire. Les valeurs des coefficients d'asymétrie et d'aplatissement sont fortement éloignées de celles qui prévalent pour une distribution gaussienne (qui par définition est symétrique et mésokurtique). La statistique de Wald proposée par DOORNIK et HANSEN [6] rejette également l'hypothèse nulle de normalité de la distribution non conditionnelle de la série en différences premières. La représentation de $(\tilde{z}_t - \tilde{z}_{t-1})$ sous la forme d'une suite de variables aléatoires indépendantes identiquement distribuées $(i.i.d)$ dont les moments du second ordre existent nécessite de tester l'existence de dépendances du type mémoire longue ou non linéaires. Les tableaux 3 et 4 présentent les autocorrélations empiriques du processus en valeur absolue, les statistiques R/S de Hurst-Mandelbrot et R/S corrigée de Lo avec 20 retards, une estimation du degré d'intégration fractionnaire par la méthode de Geweke et Porter-Hudak (GPH) pour différentes valeurs n^k d'ordonnées spectrales, et la statistique portmanteau de non linéarité de Brock, Dechert et Sheinkman (BDS(m,ε)), avec différentes valeurs de m et ε représentant respectivement la dimension de plongement et le facteur multiplicatif de l'écart type des données. Le test BDS présente l'avantage d'être applicable dans le cas où la suite de variables aléatoires ne possède pas de moments du second ordre (le test est appliqué ici directement sur les données en différences premières).

Les accroissements de taux d'intérêt présentent un grand nombre de dépendances non linéaires. Les statistiques BDS sont significatives pour l'ensemble des valeurs retenues pour le couple (m,ε). Les fonctions d'autocorrélations en valeur absolues sont très élevées et décroissent faiblement avec le nombre de retards, favorisant l'introduction de non linéarités sous la forme d'hétéroscédasticité conditionnelle. La représentation du taux d'intérêt comme un processus stationnaire en différences premières est confirmée par l'examen des tests de mémoire longue. Toute représentation de la densité conditionnelle $p(z_t \mid x_t, \rho) = p(z_t \mid z_{t-L}, \ldots, z_{t-1}, \rho)$ se doit donc d'intégrer l'ensemble de ces particularités.

3.2 Représentation du modèle auxiliaire

Le modèle auxiliaire ou générateur de scores retenu ici est l'estimateur de densité non paramétrique de GALLANT et NYCHKA [9]. Celui-ci est choisi parmi une classe de densités conditionnelles de la forme

$$\mathcal{D}_K = \{f_K(z_t \mid z_{t-L}, \ldots, z_{t-1}, \theta) : \theta = (\theta_1, \ldots, \theta_{p_K})\} \tag{3.1}$$

avec

$$f(z_t \mid z_{t-L}, \ldots, z_{t-1}, \theta) \in \mathcal{D} = \bigcup_{K=1}^{\infty} \{f_K(z_t \mid z_{t-L}, \ldots, z_{t-1}, \theta)\}$$

Si $\tilde{\theta}$ représente l'estimateur du pseudo maximum de vraisemblance de θ et si K croît avec la taille de l'échantillon n alors

$$\tilde{p}(z_t \mid z_{t-L}, \ldots, z_{t-1}, \rho) = f(z_t \mid z_{t-L}, \ldots, z_{t-1}, \tilde{\theta}) \tag{3.2}$$

est un estimateur de $p(z_t \mid z_{t-L}, \ldots, z_{t-1}, \rho)$ non paramétrique, convergent, et efficace, doté de bonnes propriétés (FENTON et GALLANT [8]). La densité conditionnelle est alors représentée par

$$f(z_t \mid x_t, \theta) \propto [P_K(y_t, x_t)]^2 \phi(y_t)[H_{x_t}]^{-\frac{1}{2}} \tag{3.3}$$

où $P_K(y_t, x_t) = \Sigma_{i=0}^{K_y} \Sigma_{j=0}^{K_x} e_{ji} x_t^j y_t^i$ est un polynôme de degré K_y en y_t et dont les coefficients sont des polynômes de degré K_x en x_t avec $e_{00} = 1$ et L_x retards de x_t, $\phi(\cdot)$ représente la densité de la loi normale centrée réduite, $y_t = H_{x_t}^{-1/2}(z_t - \mu_{x_t})$ est une innovation *i.i.d*, μ_x est une fonction de valeur centrale ou de moyenne conditionnelle autorégressive (AR)

$$\mu_{x_t} = \alpha + \beta x_t$$

qui dépend de $L_\mu \leq L$ retards avec α et β des vecteurs de paramètres, et où H_{x_t} représente une fonction de dispersion ou de variance conditionnelle hétéroscédastique (ARCH) représentée par

$$H_{x_t} = \delta_0 + \delta | x_t - \mu_{x_{t-1}} |$$

qui dépend de L_h innovations non standardisées retardées $(z_t - \mu_{x_t})$ en valeurs absolues et de $(L_h + L_\mu) \leq L$ z_t retardés, et où δ est un vecteur de paramètres. Le terme $\left[P_K(y_t, x_t)\right]^2 \phi(y)$ s'interprète comme un polynôme d'Hermite, et la constante de proportionalité $\int \left[P_K(u, x_t)\right]^2 \phi(u) \mathrm{d}(u)$ fait $f(z_t \mid x_t, \theta)$ s'intégrer à l'unité. Sous cette représentation, le polynôme d'Hermite ajuste l'ensemble des caractéristiques non gaussiennes de la distribution conditionnelle. Il s'acquitte d'autant mieux de cette tâche que les deux premiers moments conditionnels sont correctement approximés. Le tableau 5 présente les différentes caractéristiques représentées par $f(z_t \mid x_t, \theta)$ selon la nature des contraintes posées sur les paramètres L_μ, L_h, L_x, K_y, et K_x. Ces derniers sont déterminés lors de l'étape d'estimation à l'aide des critères d'information d'Akaike (AIC), d'Hannan-Quinn (HQ), et de Schwarz (BIC), et par l'examen des propriétés statistiques des résidus standardisés en utilisant la statistique BDS. Contrairement au cas de résidus de modèles linéaires, la distribution asymptotique de la statistique BDS est affectée lors de son application aux résidus de modèles à structure ARCH. Un ensemble de simulations Monte Carlo est alors réalisé pour les modèles présentant les meilleurs résultats au regard des critères d'information utilisés afin de déterminer des seuils critiques corrigés.

3.3 Estimation du modèle auxiliaire

Le tableau 6 présente le résultat des estimations par pseudo maximum de vraisemblance et détaille pour chaque meilleur modèle appartenant à l'une des classes présentées dans le tableau 5 la valeur maximisée de la fonction de log-vraisemblance (LV) ainsi que les valeurs obtenues pour les différents critères d'information selon le nombre de paramètres introduits dans le développement des fonctions de moments conditionnels et du polynôme d'Hermite[4]. Pour pouvoir comparer l'ensemble des modèles, les 15 premières observations sont réservées à la construction des retards.

Le degré élevé de persistance de la série de taux d'intérêt conjugué à une série en différences premières assimilable à un bruit blanc faible nous incite à retenir un modèle autorégressif d'ordre un sur la série de taux d'intérêt. Pour l'ensemble des spécifications testées, le polynôme AR présente une quasi racine unité (la valeur estimée du paramètre autorégressif est égale à 0,99745 pour le générateur de scores retenu pour l'estimation du modèle structurel, avec un écart-type de 0,00163). La longueur de l'échantillon permet cependant de rejeter l'hypothèse de non stationnarité du fait d'un faible écart-type, et donc de favoriser la stabilité dynamique du modèle auxiliaire en moyenne conditionnelle. La seconde étape du processus d'estimation du modèle auxiliaire consiste à déterminer la représentation conditionnellement hétéroscédastique optimale. Un nombre de retards $L_h = 7$ est sélectionné, donnant lieu à un modèle ARCH gaussien (ARCH-G). Le terme non paramétrique de la distribution conditionnelle est alors représenté en ajustant les paramètres K_y et K_x. Un premier modèle ARCH non gaussien, noté ARCH-NG, est identifié, avec une valeur $K_y = 8$. Le degré élevé du polynôme d'Hermite reflète la présence de queues de distribution conditionnelle épaisses. L'incrément en $K_x = 1$ introduit une hétérogénéité dans la distribution conditionnelle et donne lieu à un modèle non linéaire non gaussien (NL-NG). Les critères d'Akaike et d'Hannan-Quinn obtenus sont inférieurs à ceux du modèle ARCH-NG, mais le critère de Schwarz, plus restrictif sélectionne le modèle ARCH-NG. Le choix du générateur de scores se fait donc entre ces deux modèles. Le critère retenu est celui de stabilité dynamique, condition nécessaire à la mise en œuvre des méthodes d'estimation par simulation du modèle structurel[5]. Pour cela, de longues simulations (100 000 observations) et 1 000 simulations de la longueur de l'échantillon sont effectuées pour chaque modèle. A ce stade, le modèle ARCH-NG de paramètres $L_\mu = 1$, $L_h = 7$, et $K_y = 8$, s'avère très satisfaisant. Les figures 2 et 3 présentent des estimations non paramétriques par

4. Les estimations des modèles auxiliaires et structurels ont été réalisées sur PC Pentium Pro et stations de travail UltraSparc, avec des variations marginales selon les plates-formes reflétant les différents compilateurs utilisés. Nous remercions Ronald Gallant et George Tauchen pour la mise à notre disposition de leurs programmes à des fins de validation croisée.

5. Au-delà des contraintes liées à l'estimation, la justification économique d'un modèle explosif pour représenter les propriétés statistiques du taux d'intérêt à court terme est difficilement envisageable.

noyau gaussien des distributions du maximum et du minimum sur l'ensemble des 1 000 simulations. Le modèle ARCH-NG présente des caractéristiques de stationnarité et est éligible comme générateur de scores[6]. Pour conclure, le test BDS est appliqué aux résidus standardisés du modèle ARCH-NG afin de tester la présence de non linéarités résiduelles. Le tableau 7 présente les résultats obtenus. Les valeurs des statistiques sont généralement près de 10 fois inférieures à celles calculées sur la série en différences premières, avec cependant certaines statistiques significatives aux seuils critiques simulés à 5 % pour des valeurs élevées de ε, laissant supposer quelques lacunes de modélisation. La présence assez importante de chocs exceptionnels dans la série de taux d'intérêt que notre modèle ne peut expliquer contribue certainement à ces résultats. Nous retenons néanmoins le modèle ARCH-NG comme générateur de scores pour l'estimation MME.

4. Estimation de diffusions

4.1 Choix de modèles et résultats

Quatre différentes spécifications sont retenues pour l'estimation d'un modèle structurel représentatif du taux d'intérêt à court terme français : deux spécifications à un facteur communément utilisées en économie financière (les modèles MBG-plus et MBG-plus $_{\gamma\ libre}$), et deux modèles à deux facteurs(DVS et DVS $_{\gamma\ libre}$). L'évaluation de l'intégrale (2.5) est effectuée sur la base d'un échantillon final de 75 000 observations simulées $\big(S(N+L) = 1\,050\,210\big)$ déterminé à partir de 5 000 premières réalisations L. Les résultats obtenus sont stables pour différentes valeurs de $S \geq 7$ et $N \geq 50\,000$. Pour chaque paramètre des spécifications en temps continu, les estimations ainsi que les écarts-types de Wald sont présentés. L'estimation MME présente alors des valeurs estimées de paramètres directement interprétables.

4.2 Résultats

Le tableau 8 récapitule les résultats du modèle MBG-plus. Ce modèle introduit plus de flexibilité dans le traitement de notre spécification à un facteur du fait d'une ordonnée à l'origine non nulle pour la fonction de diffusion. Le modèle est rejeté aux seuils statistiques usuels. L'étude des scores souligne l'apparente justification du modèle structurel pour la représentation de la

6. L'absence de contraintes de non négativité lors de l'estimation du modèle auxiliaire conduit à une distribution empirique du minimum de chaque sentier simulé présentant des valeurs négatives. L'intégration d'un processus du type racine carrée dans la dynamique de volatilité permet de résoudre ce problème. La question fondamentale pour l'estimation MME demeure néanmoins la stabilité dynamique du modèle auxiliaire, condition assurée ici.

dynamique de moyenne. L'état stationnaire pour le taux d'intérêt $(-t_{11}/t_{12})$ évalué à 3,64 % est néanmoins différent de la moyenne de l'échantillon. Un intervalle de confiance à 95 % sur le paramètre $-t_{11}/t_{12}$ s'évalue à (1,45 %, 9,67 %), et confirme la difficile estimation des paramètres de tendance. La valeur du paramètre t_{12} est strictement négative et garantit ainsi l'absence de racine unitaire dans la dynamique de moyenne. La fonction de diffusion se dissipe en 3,32 %, générant une barrière réfléchissante à droite. Les ratios de Student relatifs aux scores décrivant la fonction de variance conditionnelle et le polynôme d'Hermite sont également élevés et dénotent une mauvaise spécification de la dynamique de volatilité et la difficulté du modèle à un facteur à représenter le caractère non gaussien de l'échantillon.

Le tableau 9 (modèle MBG-plus $_{\gamma\ libre}$) présente l'estimation du modèle à un facteur augmenté du paramètre γ. Les résultats sont similaires à ceux obtenus précédemment, avec un paramètre γ non significativement différent de l'unité. La procédure d'optimisation numérique est néanmoins particulièrement sensible au choix des valeurs initiales pour ce paramètre.

Les limitations des modèles à un facteur semblent provenir de deux sources de mauvaise spécification : incapacité à représenter les propriétés non gaussiennes de l'échantillon d'une part, et absence d'une quelconque dynamique aléatoire de volatilité d'autre part. Afin de corriger les modèles de taux d'intérêt estimés de ces potentielles mauvaises spécifications, les représentations à un facteur sont augmentées d'une composante aléatoire non observable.

Les résultats obtenus pour les spécifications DVS et DVS $_{\gamma\ libre}$ sont exposés dans les tableaux 10 et 11. Les moyennes non conditionnelles des deux processus sont respectivement de 3,346 % et de 3,406 %. Le modèle DVS améliore très sensiblement les modèles MBG-plus et MBG-plus $_{\gamma\ libre}$. L'introduction d'un second facteur relatif à la dynamique de volatilité, suggéré par l'analyse des scores individuels des modèles à un facteur, s'avère une réussite. Ainsi, les scores représentatifs du second moment conditionnel sont non significatifs. De plus, les caractéristiques non gaussiennes de l'échantillon sont presque entièrement capturées comme l'atteste la faible significativité des scores relatifs au polynôme d'Hermite. La statistique du khi-deux de bonne spécification globale du modèle structurel rejette largement l'hypothèse alternative. La diffusion DVS semble représenter l'ensemble des caractéristiques de la densité conditionnelle de l'échantillon.

Enfin, l'estimation du modèle DVS $_{\gamma\ libre}$ atteste de la nécessité d'incorporer une spécification aléatoire pour le processus de volatilité. Le modèle améliore les résultats obtenus avec la représentation DVS en termes de significativité globale et de la fonction score du modèle auxiliaire. La valeur obtenue pour le paramètre γ est beaucoup plus faible que dans le cas du modèle MBG-plus $_{\gamma\ libre}$, et significativement non différente de 0,5. La représentation racine carrée pour le processus de taux d'intérêt semble être valide dans le cadre de modèles multifacteurs qui permettent de discriminer entre hétéroscédasticité conditionnelle et lien entre niveau et volatilité du taux d'intérêt instantané.

5. Conclusion

Cette recherche applique une classe d'estimateurs du khi-deux minimum à l'estimation de systèmes d'équations différentielles stochastiques en temps continu représentant un modèle structurel. La méthode des moments efficaces, dont les propriétés sont, sous certaines conditions, similaires à la méthode du maximum de vraisemblance, permet une estimation directe des paramètres d'une diffusion dans le cas où les densités jointe et de transition ne possèdent pas de forme dérivable analytiquement.

Les modèles à un facteur ne permettent pas de représenter la dynamique du taux d'intérêt à court terme français. L'introduction d'un second facteur non observable relatif à la dynamique de volatilité permet d'améliorer remarquablement la cohérence entre modèle structurel et densité conditionnelle estimée représentative de la série chronologique à temps discret de taux d'intérêt.

Le succès de cette procédure d'estimation en deux étapes est cependant subordonné au choix du modèle auxiliaire. Une plus grande flexibilité de la forme fonctionnelle de ce modèle, en particulier l'introduction de différents régimes ou de seuils dans la modélisation de la moyenne conditionnelle de séries financières, apparaît comme un développement possible de l'estimation de modèles auxiliaires afin de totalement capturer les non linéarités résiduelles.

L'amélioration des méthodes d'intégration afin de réduire l'erreur Monte Carlo (en utilisant des générateurs de nombres quasi-aléatoires ou des techniques de réduction de variance) est également une possibilité de développement intéressante de la méthode des moments efficaces. De plus, la spécification du modèle structurel peut être augmentée de termes traitant plus spécifiquement du problème de la persistance propre aux séries de taux d'intérêt et de l'occurrence de grands chocs. On peut alors penser à une décomposition en composantes stationnaires et permanentes de l'évolution moyenne du processus de diffusion, ainsi qu'à l'introduction de processus à sauts de Poisson. Enfin, la recherche d'une plus grande cohérence entre modèles structurels et propriétés des séries chronologiques financières à temps discret devrait offrir une plus grande flexibilité dans la modélisation de la dynamique de la structure par terme des taux d'intérêt.

Annexes

Tableau 1 : Contraintes sur les paramètres du modèle DVS

Modèle	t_{11}	t_{12}	v_{11}	v_{12}	t_{21}	t_{22}	t_{23}	v_{21}	v_{23}	γ
Merton		0		0	0	0	0	0	0	0
Vasicek				0	0	0	0	0	0	0
Brennan-Schwartz			0		0	0	0	0	0	1
MBG-plus					0	0	0	0	0	1
MBG-plus $_{\gamma}$ *libre*					0	0	0	0	0	
DVS										1
DVS $_{\gamma}$ *libre*										

Tableau 2 : Structure statistique de l'échantillon

	$\tilde{z}_t$	$(\tilde{z}_t - \tilde{z}_{t-1})$
Moyenne	9.250	−0.009
Médiane	9.047	0.000
Ecart-type	2.550	0.295
Asymétrie	0.448	1.447
$p-$valeur	0.000	0.000
Excès de kurtosis	0.192	15.239
$p-$valeur	0.179	0.000
Maximum	17.919	2.653
Minimum	3.623	−1.699
Wald	42.328	948.971
$p-$valeur	0.000	0.000
Autocorrélations		
$\rho(1)$	0.990	−0.011
$\rho(2)$	0.980	0.036
$\rho(3)$	0.971	0.058
$\rho(4)$	0.961	−0.027
$\rho(5)$	0.951	0.008
LB(24)	21953.325	27.784
LB(36)	28609.897	38.402

Tableau 3 : Analyse des dépendances non linéaires

	$(\tilde{z}_t - \tilde{z}_{t-1})$
Autocorrélations en valeur absolue	
$\rho(1)$	0.346
$\rho(2)$	0.247
$\rho(3)$	0.148
$\rho(4)$	0.171
$\rho(5)$	0.117
LB(24)	446.871
LB(36)	461.483
R/S	1.416
R/S$_{Lo}$	1.282
GPH($k = 0.5$)	−0.040
	(0.130)
	[0.166]
GPH($k = 0.525$)	0.032
	(0.118)
	[0.165]
GPH($k = 0.55$)	0.053
	(0.106)
	[0.141]

$(\cdot)$ et $[\cdot]$ sont les écarts-types asymptotiques et empiriques estimés.

Un intervalle de confiance à 99 % pour R/S et R/S$_{Lo}$ est $(0.721, 2.098)$.

Tableau 4 : Statistiques BDS, $(\tilde{z}_t - \tilde{z}_{t-1})$

m	ε			
	0.5	1.0	1.5	2.0
2	10.262*	10.690*	10.258*	10.023*
3	12.950*	13.184*	12.343*	11.604*
4	14.458*	13.853*	12.849*	11.642*
5	16.705*	14.631*	13.417*	11.735*
6	19.675*	15.647*	13.845*	11.753*

* Statistique significative au seuil critique à 5 %.

Tableau 5 : Contraintes sur les paramètres de $f(z_t \mid x_t, \theta)$

z_t	L_μ	L_h	K_y	K_x	L_x
i.i.d. Gaussien	0	0	0	0	
AR gaussien (AR-G)		0	0	0	
AR non gaussien (AR-NG)		0		0	
ARCH gaussien (ARCH-G)			0	0	
ARCH non gaussien (ARCH-NG)				0	
Non linéaire non gaussien(NL-NG)					

Tableau 6 : Représentations de $f(z_t \mid x_t, \theta)$

z_t	LV	AIC	HQ	BIC	L_μ	L_h	K_y	K_x	L_x
AR-G	878.124	−0.743	−0.741	−0.736	1	0	0	0	0
ARCH-G	1128.126	−0.942	−0.934	−0.920	1	7	0	0	0
* ARCH-NG	1350.824	−1.131	−1.117	−1.093	1	7	8	0	0
NL-NG	1369.158	−1.1393	−1.117	−1.081	1	7	8	1	0

* Générateur de scores retenu pour l'estimation MME.

Tableau 7 : Statistiques BDS, modèle ARCH-NG

m	ε			
	0.5	1.0	1.5	2.0
2	0.436	0.095	0.721	2.112*
3	0.054	0.272	1.381	3.066*
4	0.252	0.322	1.396	3.090*
5	0.445	0.473	1.474*	3.222*
6	0.244	0.375	1.376*	3.211*

* Statistique significative au seuil critique simulé à 5 %.

Tableau 8 : Estimation de diffusion, modèle de MBG-plus

Modèle structurel	Estimation	Ecart-type	Student
t_{11}	1.601	0.421	3.805*
t_{12}	−0.440	0.115	−3.810*
v_{11}	0.744	0.135	5.493*
v_{12}	−0.224	0.031	−7.263*
Critère MMG			
$nm_N(\rho,\tilde{\theta})'(\tilde{\mathcal{I}}_n^{-1})m_N(\rho,\tilde{\theta})$	26.269		
$p_\theta - p_\rho$	14		
Modèle auxiliaire	Score moyen	Ecart-type	Student
α	11.647	17.747	0.656
β_1	−8.181	8.786	−0.931
δ_0	14.920	12.500	1.194
δ_1	−0.862	0.709	−1.216
δ_2	0.372	0.790	0.471
δ_3	−0.455	0.717	−0.635
δ_4	1.719	0.928	1.851
δ_5	2.276	0.976	2.333*
δ_6	1.794	1.046	1.745
δ_7	2.356	1.901	1.239
e_{10}	−1.922	2.433	−0.790
e_{20}	3.609	2.637	1.369
e_{30}	−12.066	9.607	−1.256
e_{40}	53.541	21.431	2.498*
e_{50}	−66.780	79.278	−0.842
e_{60}	715.059	239.282	2.988*
e_{70}	−336.948	918.496	−0.367
e_{80}	11110.402	3236.623	3.433*

Ecarts-types robustes calculés avec le jacobien $\widehat{M}.t_{21} = -t_{22}$ pour identification.

* Statistique significative au seuil critique à 5 %.

Tableau 9 : Estimation de diffusion, modèle MBG-plus $_{\gamma}$ *libre*

Modèle structurel	Estimation	Ecart-type	Student
t_{11}	1.595	0.423	3.766*
t_{12}	−0.438	0.115	−3.779*
v_{11}	0.725	0.313	2.314*
v_{12}	−0.217	0.092	−2.362*
γ	0.982	0.259	3.790*
Critère MMG			
$nm_N(\rho,\tilde{\theta})'(\tilde{\mathcal{I}}_n^{-1})m_N(\rho,\tilde{\theta})$	26.261		
$p_\theta - p_\rho$	13		
Modèle auxiliaire	Score moyen	Ecart-type	Student
α	11.688	17.715	0.660
β_1	−8.261	8.410	−0.982
δ_0	15.356	10.441	1.471
δ_1	−0.861	0.698	−1.232
δ_2	0.371	0.783	0.474
δ_3	−0.446	0.619	−0.720
δ_4	1.730	0.901	1.920
δ_5	2.291	0.946	2.420*
δ_6	1.809	1.022	1.768
δ_7	2.372	1.890	1.255
e_{10}	−1.940	2.425	−0.800
e_{20}	3.448	1.742	1.979*
e_{30}	−12.213	9.458	−1.291
e_{40}	52.542	18.365	2.861*
e_{50}	−67.910	78.463	−0.866
e_{60}	708.009	228.693	3.096*
e_{70}	−347.159	915.040	−0.379
e_{80}	11052.306	3199.651	3.454*

Ecarts-types robustes calculés avec le jacobien $\widehat{M}.t_{21} = -t_{22}$ pour identification.

* Statistique significative au seuil critique à 5 %.

Tableau 10 : Estimation de diffusion, modèle DVS

Modèle structurel	Estimation	Ecart-type	Student
t_{11}	0.722	0.268	2.694*
t_{12}	−0.216	0.042	−5.140*
v_{11}	0.172	0.036	4.810*
v_{12}	−0.061	0.013	−4.812*
t_{22}	−2.272	0.404	−5.627*
v_{21}	2.939	0.961	3.060*
v_{22}	−0.735	0.251	2.929*
Critère MMG			
$nm_N(\rho,\tilde{\theta})'(\tilde{\mathcal{I}}_n^{-1})m_N(\rho,\tilde{\theta})$	17.159		
$p_\theta - p_\rho$	11		
Modèle auxiliaire	Score moyen	Ecart-type	Student
α	8.802	19.038	0.462
β_1	−2.359	19.210	−0.123
δ_0	14.982	17.926	0.836
δ_1	0.276	0.888	0.311
δ_2	0.811	0.964	0.841
δ_3	−1.076	1.320	−0.815
δ_4	0.959	1.112	0.863
δ_5	1.987	1.079	1.840
δ_6	1.051	1.184	0.888
δ_7	1.765	1.951	0.905
e_{10}	1.727	2.523	0.684
e_{20}	3.291	3.900	0.844
e_{30}	2.035	9.781	0.208
e_{40}	44.710	25.645	1.743
e_{50}	−7.159	81.159	−0.088
e_{60}	627.579	264.384	2.374*
e_{70}	−28.809	948.728	−0.030
e_{80}	10041.381	3487.244	2.879*

Ecarts-types robustes calculés avec le jacobien $\widehat{M}.t_{21} = -t_{22}$ pour identification.

*** Statistique significative au seuil critique à 5 %.**

Tableau 11 : Estimation de diffusion, modèle DVS γ *libre*

Modèle structurel	Estimation	Ecart-type	Student
t_{11}	0.786	0.287	2.738*
t_{12}	−0.231	0.155	−1.490*
v_{11}	0.0648	0.0285	2.272*
v_{12}	−0.021	0.009	−2.198*
t_{22}	−3.134	0.477	−6.572*
v_{21}	3.066	0.718	4.270*
v_{22}	−0.642	0.216	−2.976*
γ	0.681	0.253	−2.692*
Critère MMG			
$nm_N(\rho,\tilde{\theta})'(\tilde{\mathcal{I}}_n^{-1})m_N(\rho,\tilde{\theta})$	15.908		
$p_\theta - p_\rho$	10		
Modèle auxiliaire	Score moyen	Ecart-type	Student
α	8.661	19.038	0.455
β_1	−1.806	19.210	−0.94
δ_0	15.000	17.926	0.837
δ_1	0.090	0.888	0.102
δ_2	0.527	0.964	0.547
δ_3	−1.233	1.321	−0.934
δ_4	0.844	1.111	0.759
δ_5	1.979	1.079	1.833
δ_6	0.963	1.184	0.814
δ_7	1.849	1.951	0.948
e_{10}	1.206	2.523	0.478
e_{20}	1.885	3.900	0.483
e_{30}	−0.725	9.781	−0.074
e_{40}	33.126	25.646	1.292
e_{50}	−23.832	81.163	−0.294
e_{60}	518.166	364.400	1.422
e_{70}	−159.134	945.788	−0.168
e_{80}	8764.812	3487.462	2.513*

Ecarts-types robustes calculés avec le jacobien $\widehat{M}.t_{21} = -t_{22}$ pour identification.

*** Statistique significative au seuil critique à 5 %.**

Références

[1] M. BRENNAN et E. SCHWARTZ (1980) "Analyzing Convertible Bonds", *Journal of Financial and Quantitative Analysis,* 15 : 907-929.

[2] R. J. BRENNER, R. H. HARJES, et K. F. KRONER (1996) "Another Look at Models of the Short-term Interest Rate", *Journal of Financial and Quantitative Analysis,* 31 : 85-107.

[3] L. BROZE, O. SCAILLET, et J.M. ZAKOIAN (1993) *Testing for Continuous-time Models of the Short-term Interest Rate,* Unpublished Manuscript, CORE, Belgium.

[4] K. C. CHAN, G. A. KAROLYI, F. A. LONGSTAFF, et A. B. SANDERS (1992) "An Empirical Comparison of Alternative Models of the Short-term Interest Rate", *Journal of Finance,* 47 : 1209-1227.

[5] J. COX, J. INGERSOLL, et S. ROSS (1985) "A Theory of the Term Structure of Interest Rates", *Econometrica,* 53 : 145-166.

[6] J. A. DOORNIK et H. HANSEN (1994) *An Omnibus Test for Univariate and Multivariate Normality,* Unpublished Manuscript, Nuffield College, Oxford.

[7] B. P. FLANNERY, W. H. PRESS, S. A. TEUKOLSKY, et W. T. VETTERLING (1994) *Numerical Recipes in Fortran : the Art of Scientific Computing,* Second edition, Cambridge University Press, New York.

[8] A. R. GALLANT et V. M. FENTON (1996) "Convergence Rates of SNP Density Estimators", *Econometrica,* 64 : 719-727.

[9] A. R. GALLANT et D. W. NYCHKA (1987) "Semi-nonparametric Maximum Likelihood Estimation", *Econometrica,* 55 : 363-390.

[10] A. R. GALLANT et G. TAUCHEN (1996) "Which Moments to Match ?", *Econometric Theory,* 12 : 657-681.

[11] A. J. KINDERMANN et J. F. MONAHAN (1977) "Computer Generation of Random Variables Using the Ratio of Uniform Deviates", *ACM Transactions on Mathematical Software* 3 : 257-260.

[12] P. E. KLOEDEN et E. PLATEN (1992) *Numerical Solution of Stochastic Differential Equations,* Springer-Verlag, New York.

[13] R. MERTON (1973) "Theory of Rational Option Pricing", *Bell Journal of Economics and Management Science,* 4 : 141-183.

[14] G. TAUCHEN (1997) "The Objective Function of Simulation Estimators near the Boundary of the Stability Set", Forthcoming *The Review of Economics and Statistics.*

[15] O. VASICEK (1977) "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics,* 5 : 177-188.