JOURNAL DE LA SOCIÉTÉ FRANÇAISE DE STATISTIQUE

ÉRIC RENAULT

La mémoire longue en économie : discussion et commentaires

Journal de la société française de statistique, tome 140, nº 2 (1999), p. 79-89

<http://www.numdam.org/item?id=JSFS_1999__140_2_79_0>

LA MÉMOIRE LONGUE EN ÉCONOMIE : DISCUSSION ET COMMENTAIRES

Eric RENAULT *

L'article discuté ici (désigné ci-après par LM) donne à juste titre deux directions d'application de la mémoire longue en économie : les séries temporelles macroéconomiques et les séries financières, qu'il s'agisse de séries de prix d'actifs financiers (actions, taux de change, taux d'intérêt) ou du processus de la variance conditionnelle d'un rendement (notion de volatilité en finance). Cette discussion sera concentrée sur la mise en exergue de certains écueils de la modélisation à mémoire longue en finance et de certains outils pour y faire face. C'est pourquoi, plutôt que de revenir sur les multiples éclairages intéressants fournis par LM, on essayera de fournir des compléments sur quatre points, tous relatifs à la modélisation de la dynamique de la volatilité. On n'oubliera évidemment pas qu'un but central de la finance est la modélisation des prix d'actifs financiers mais on soutiendra (voir notre quatrième point) que la mémoire longue est peu défendable comme outil de prévision des rendements eux-mêmes, mais bien davantage comme modèle de la dynamique de leur volatilité. Cela n'est pas sans conséquence pour l'évaluation des actifs financiers, mais plus que des actifs primitifs (action, intérêt, change) il s'agit de l'évaluation des actifs dérivés (options) construits sur des sous-jacents dont la volatilité est à mémoire longue. Les thèses que nous développons sont pour l'essentiel non originales mais inspirées d'articles publiés dont plusieurs ont échappé à la revue de la littérature, par ailleurs très complète, de LM. Il nous a semblé naturel de donner d'abord la référence de ces articles en regard de l'intitulé de chacun des quatre points afférents, avant de détailler davantage chaque argument. La discussion s'articulera sur l'enchaînement des quatre thèses suivantes :

1. Les processus FIGARCH (section 4.2. de LM) ne peuvent pas fournir un outil tel que « par analogie avec les processus ARFIMA, la dynamique de long terme de la volatilité est (serait) prise en compte au travers du paramètre d'intégration fractionnaire d ». En fait, ces processus ne peuvent pas décrire une dynamique stationnaire en dehors du cas $d = 0$, c'est-à-dire du modèle GARCH à mémoire courte usuel.

* CREST-ENSAI, Campus de Ker Lann, 35170 Bruz et Université de Montréal
e-mail : erenault@ensai.fr

Journal de la Société Française de Statistique, tome 140, n° 2, 1999

Références complémentaires :

GHYSELS E., HARVEY A.C. et RENAULT E. (1996) : "Stochastic Volatility" in Handbook of Statistics, Vol 14, p 119-191, G.S. Maddala et C.R. Rao éditeurs.

ROBINSON P.M. (1991) : "Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression", Journal of Econometrics, 47, p 67-84.

ROBINSON P.M. et ZAFFARONI P. (1997) : "Modelling Nonlinearity and Long Memory in Time Series", Fields Institute Communications, Vol 11, p 161-170.

2. Si, comme souligné dans LM (section 2.2.2.), «les premiers processus à mémoire longue ont été développés en temps continu», il est regrettable que ceci ait été oublié ensuite car il faudrait considérer une paramètrisation des processus ARFIMA définie en temps continu (représentations AR ou MA par rapport au brownien fractionnaire) pour pouvoir affirmer à juste titre qu'ils «présentent l'intérêt de tenir compte à la fois du comportement de court terme de la série au travers des paramètres autorégressifs et moyenne mobile et du comportement de long terme par le biais du paramètre d'intégration fractionnaire». En effet, quand un tel processus est échantillonné en temps discret, on obtient effectivement un processus ARFIMA mais dont les coefficients AR et MA dépendent de façon compliquée du paramètre d d'intégration fractionnaire.

   Références complémentaires :

   COMTE F. (1996) : "Simulation and Estimation of Long Memory Continuous Time Models", Journal of Time Series Analysis,17, p 19-36.

   COMTE F. et RENAULT E. (1996) : "Long Memory Continuous Time Models", Journal of Econometrics, 73, p 101-149.

3. Pour avoir des modèles de la dynamique de la volatilité en temps discret qui soient compatibles avec un modèle en temps continu sous-jacent, il est nécessaire de relâcher la contrainte de la structure de type GARCH dans laquelle la variance conditionnelle est représentée par une fonction déterministe des innovations passées. On définit ainsi un modèle dit à volatilité stochastique, qui entre dans la classe des modèles à chaîne de Markov cachée. Ce modèle est robuste à l'agrégation temporelle (la structure du modèle ne change pas lorsque l'on change la fréquence d'observation) alors que la contrainte sur ce modèle qui définit la sous-classe des processus GARCH ne peut pas être valide à deux fréquences différentes.

   Références complémentaires :

   DROST F.C. et NIJMAN Th. E. (1993) : "Temporal Aggregation of GARCH Processes", Econometrica,61, p 909-927.

   DROST F.C. et WERKER B.J.M. (1996) : "Closing the GARCH gap : Continuous Time GARCH Modeling", Journal of Econometrics, 74, p 31-58.

MEDDAHI N. et RENAULT E. (1996) : "Aggregation and Marginalization of GARCH and Stochastic Volatility Models", GREMAQ DP 96.30.433, Université des Sciences Sociales de Toulouse.

4. La modélisation à mémoire longue en Finance apparaît, autant pour des raisons empiriques que théoriques, surtout pertinente pour les processus de volatilité. Elle permet alors d'améliorer significativement la performance empirique des modèles d'évaluation d'option, particulièrement dans le cas d'options de maturité longue. On obtient des modèles qui généralisent les modèles usuels (formules d'évaluation de Black-Scholes (1973) et de Hull-White (1987)) grâce à la définition de processus de volatilité stochastique à mémoire longue en temps continu.

Références complémentaires :

BOLLERSLEV T. et MIKKELSEN H.O. (1999) : "Long Term Equity Anticipation Securities and Stock Market Volatility Dynamics", Journal of Econometrics, 92, p 75-100.

COMTE F. et RENAULT E. (1998) : "Long Memory in Continuous Time Stochastic Volatility Models", Mathematical Finance, 8, p 291-323.

RENAULT E. (1997) : "Econometric models of Option Pricing Errors" in Advances in Econometrics, Seventh World Congress , Vol.3 (eds. D.M. Kreps et K.F. Wallis), Cambridge University Press.

RENAULT E. et TOUZI N. (1996) : "Option Hedging and Implicit Volatilities in a Stochastic Volatility Model", Mathematical Finance, 6, 279-302.

1. LES PROCESSUS FIGARCH NE SONT PAS STATIONNAIRES

Comme souligné en introduction de LM, une motivation essentielle de la modélisation à mémoire longue est de construire des processus qui sont « caractérisés par une fonction d'autocorrélation décroissant hyperboliquement ». Les processus ARFIMA remplissent cette condition mais pas les processus FIGARCH car « ils » (ou plus exactement leur variance conditionnelle dont on voudrait caractériser la dynamique) n'ont pas de fonction d'autocorrélation ! Plus précisément, l'équation (86) de l'article :

$$\Phi(L)(1-L)^d\varepsilon_t^2 = \alpha_0 + [1-\beta(L)]u_t$$

ne permet que fallacieusement de retrouver « clairement l'analogie entre les processus FIGARCH sur l'équation de la variance conditionnelle et les processus ARFIMA pour l'équation de la moyenne », c'est-à-dire l'équation (8) de LM :

$$\Phi(L)(1-L)^d X_t = \Theta(L)u_t$$

Une différence essentielle entre les deux équations est en effet la présence du terme α_0 qui, s'il est non nul, interdit au processus ε_t^2 d'être stationnaire au second ordre. D'un autre côté, si ce terme est nul, la variance inconditionnelle (la valeur limite de la variance conditionnelle) dans le processus GARCH défini par (79) est elle-même nulle. Autrement dit, on ne peut en aucun cas voir les processus FIGARCH comme « un cas intermédiaire entre les processus GARCH et les processus IGARCH ».

Robinson et Zaffaroni (1997) expliquent clairement le problème et ce qu'il convient de faire pour y remédier. La solution se trouvait en fait déjà dans Robinson (1991) à qui on devrait attribuer la paternité des processus GARCH à mémoire longue. En suivant leurs notations, on écrira donc la représentation AR∞ du processus stationnaire au second ordre (on ne veut pas sortir de ce cadre) ε_t^2 :

$$\varepsilon_t^2 = \sigma^2 + \sum_{j=1}^{\infty} \psi_j [\varepsilon_{t-j}^2 - \sigma^2] + u_t$$

où $\sigma^2 = E[\varepsilon_t^2] = E[\sigma_t^2]$.

On notera qu'avant d'introduire un terme de tendance $(1-L)^d$, il convient toujours de travailler (comme on en a l'habitude pour les processus ARFIMA) sur une représentation ARMA (éventuellement infinie) du processus **centré**. En notant :

$$\psi(L) = 1 - \sum_{j=1}^{\infty} \psi_j L^j$$

la variance conditionnelle de ε_t s'écrit :

$$\sigma_t^2 = \psi(1)\sigma^2 + [1 - \psi(L)]\varepsilon_t^2$$

si bien que l'on retrouve les processus usuels :

- ARCH(p) dans le cas de coefficients ψ_j nuls pour $j > p$,
- GARCH dans le cas de coefficients ψ_j décroissant à vitesse exponentielle,
- mémoire longue (dans σ_t^2 comme dans ε_t^2) dans le cas où il existe $c > 0$ et un coefficient d, $0 < d < 1/2$, tels que :

  $$|\psi(\exp i\lambda)| \sim c\lambda^d \text{ quand } \lambda \longrightarrow 0^+$$

  de façon que la densité spectrale de ε_t^2 soit de l'ordre de λ^{-2d} quand $\lambda \longrightarrow 0^+$.

L'adaptation de l'idée des ARFIMA au contexte GARCH consiste donc à prendre :

$$\psi(L) = \Phi(L)(1-L)^d / \Theta(L)$$

ce qui donne :

$$\Phi(L)(1-L)^d \varepsilon_t^2 = \Theta(L) u_t$$

c'est-à-dire une équation du type (86) mais avec :

$$\alpha_0 = \psi(1)\sigma^2 = 0 \text{ bien que la variance inconditionnelle } \sigma^2 \text{ soit non nulle.}$$

Robinson et Zaffaroni (1997) soulignent cependant que ce modèle n'est pas très satisfaisant d'un point de vue statistique, en particulier parce que « aucune théorie de la distribution asymptotique des estimateurs du maximum de vraisemblance de ce modèle ne semble exister ». Ils privilégient une modélisation alternative basée sur l'idée de « moyenne mobile non linéaire ».

2. LES PARAMÈTRES DE LA DYNAMIQUE DE COURT TERME NE PEUVENT ÊTRE CARACTÉRISÉS (SANS INTERFÉRENCE AVEC LA DESCRIPTION EN BASSE FRÉQUENCE) QU'À TRAVERS UNE REPRÉSENTATION MOYENNE MOBILE EN TEMPS CONTINU PAR RAPPORT AU BROWNIEN FRACTIONNAIRE

Suivant Comte et Renault (1996), CR(96) dans la suite, nous nous concentrons dans ce second point sur le débat temps discret versus temps continu pour la construction des processus ARFIMA. L'application de cette construction aux processus de volatilité ne sera développée qu'au quatrième point. Pour simplifier la présentation, on se contentera ici de considérer l'analogue en temps continu des processus du type ARFIMA$(1, d, 0)$ en renvoyant à CR(96) pour une généralisation à des ARFIMA d'ordres quelconques. Il est bien connu qu'un processus AR(1) gaussien, avec autocorrélation positive, peut toujours être interprété comme l'échantillonnage en temps discret d'un processus d'Ornstein-Ulhenbeck, que l'on écrira (en se limitant au cas centré) :

$$dX(t) = -kX(t)\ dt + \sigma\ dW(t)$$

où W est un brownien standard et k et σ deux réels positifs. Avec la condition initiale $X(0) = 0$, la solution de cette EDS (Equation Différentielle Stochastique) est en effet le processus gaussien :

$$X(t) = \int_{[0,t]} \exp[-k(t-s)]\sigma\ dW(s)$$

La version « temps continu » de la construction des ARFIMA peut alors être présentée de deux façons équivalentes :

(i) Soit, de manière analogue à l'équation (8) de LM, on remplace le « bruit » $dW(t)$ par un accroissement du brownien fractionnaire défini (conformément à la formule (13) de LM) par :

$$W_d(t) = 1/\Gamma(1+d) \int_{[0,t]} (t-s)^d\ dW(s)$$

ce qui conduit à l'EDS fractionnaire :

$$dX(t) = -kX(t)\ dt + \sigma\ dW_d(t)$$

dont on montre encore que la solution s'écrit :

$$X(t) = \int_{[0,t]} \exp[-k(t-s)]\sigma \, dW_d(s)$$

étant entendu que l'intégration par rapport au brownien fractionnaire (qui n'est pas une semi-martingale) se définit pour des intégrandes C déterministes et de classe C^1 par une formule d'intégration par parties :

$$X(t) = \int_{[0,t]} C(t-s) \, dW_d(s) = 1/\Gamma(1+d) \int_{[0,t]} (t-s)^d A(t-s) dW(s)$$

où :

$$A(t) = C(0) + \int_{[0,t]} C'(s)(1-s/t)^d \, ds.$$

(ii) Soit, de manière analogue à l'équation (44) de LM, on dit que c'est la différenciation fractionnaire d'ordre d de $X(t)$ qui est le processus d'Ornstein-Ulhenbeck classique, cette différenciation fractionnaire étant définie en temps continu par l'opération inverse de celle qui fait passer du brownien standard $W(t)$ au brownien fractionnaire $W_d(t)$.

CR(96) montrent bien que les deux définitions coïncident et permettent de construire un processus $X(t)$ en temps continu, asymptotiquement stationnaire (le processus étant initialisé par la condition $X(0) = 0$), et à mémoire longue au sens où la covariance entre $X(t)$ et $X(t+h)$ est d'ordre h^{2d-1}, exactement comme dans le cas d'un processus ARFIMA(p, d, q). Le différencié (en temps continu) d'ordre d de $X(t)$ étant un processus AR(1), noté $X^{(d)}(t)$, on a d'ailleurs bien construit l'analogue d'un ARFIMA$(1, d, 0)$ en temps continu.

Mais, il est important de remarquer que la dynamique de court terme du processus d'intérêt $X(t)$ peut être très différente de celle du processus ARFIMA$(1, d, 0)$

$Y(t) = (1-L)^{-d}X^{(d)}(t)$ même s'ils tendent à être dans une relation d'équilibre de long terme puisque CR(96) montrent qu'ils sont en fait dans une relation de cointégration fractionnaire (voir section 5.1.2.2. de LM). Autrement dit, leur différence est un processus certes à mémoire courte mais en aucune façon négligeable à court terme. Pour s'en convaincre, il suffit, comme le font CR(96), d'examiner le corrélogramme partiel de $(1-L)^d X(t)$ et d'observer qu'il est très loin de s'annuler au delà du premier retard comme celui du processus AR(1)$X^{(d)}(t)$.

La conclusion de cet exercice est que la dynamique de court terme doit être caractérisée par les paramètres k et σ du processus d'Ornstein-Ulhenbeck $X^{(d)}(t)$ et non pas par les paramètres du processus $(1-L)^d X(t)$ qui sont des fonctions compliquées (il ne s'agit pas en général d'un ARMA d'ordres finis) non seulement de k et σ mais aussi de d ; on ne peut séparer les caractéristiques de court terme de celles de long terme autrement qu'en référant au modèle en temps continu sous-jacent.

3. MODÉLISATION GARCH EN TEMPS CONTINU ET VOLATILITÉ STOCHASTIQUE

La notion de GARCH intégré est présentée par LM à partir de la formulation (83), c'est-à-dire à partir de la remarque importante qu'un « processus GARCH(p,q) peut en outre s'interpréter comme un processus ARMA(m,p), avec $m = \max(p,q)$, sur le carré des innovations ». Ainsi, en posant :

$$u_t = \varepsilon_t^2 - \sigma_t^2$$

on peut transformer l'équation GARCH(1,1) (on se limite ici pour la commodité d'exposition à l'ordre un mais tout ce qui est fait dans cette section peut être généralisé à des ordres supérieurs, voir Meddahi et Renault (1996)) :

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

en : $\varepsilon_t^2 - u_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta(\varepsilon_{t-1}^2 - u_{t-1})$
c'est-à-dire :

$$(1-\gamma L)\varepsilon_t^2 = \omega + (1-\beta L)u_t \text{ avec } \gamma = \alpha + \beta.$$

Cette remarque est souvent faite pour justifier le modèle GARCH ; il serait après tout très général puisque fondé sur le théorème de Wold permettant de représenter par un ARMA (éventuellement d'ordre infini) la dynamique du processus ε_t^2 supposé stationnaire au second ordre. Cette apparence de généralité est en fait quelque peu fallacieuse puisque le théorème de Wold permettrait simplement de dire que ε_t^2 est un processus ARMA avec un processus d'innovations u_t qui est un bruit blanc au second ordre, c'est-à-dire sans corrélation sérielle. Alors, l'équation :

$$\varepsilon_t^2 = \omega + \gamma\varepsilon_{t-1}^2 - \beta u_{t-1} + u_t \qquad (*)$$

s'interpréterait simplement en disant que $\omega + \gamma\varepsilon_{t-1}^2 - \beta u_{t-1}$ est la projection orthogonale de ε_t^2 sur l'espace de Hilbert engendré par les u_s, $s < t$. C'est fondamentalement (nous négligeons volontairement ici quelques difficultés techniques) la notion de **GARCH faible** introduite par Drost et Nijman (1993). La notion de GARCH usuelle, utilisée par LM, est plus restrictive puisqu'elle suppose en outre que u_t est une différence de martingale :

$$E[u_t/u_s, s < t] = 0$$

Suivant la terminologie introduite par Drost et Nijman (1993), on parlera dans ce cas de **GARCH semi fort** (on ne parlera de **GARCH fort** qu'avec l'hypothèse additionnelle que les innovations standardisées ε_t/σ_t sont indépendantes, identiquement distribuées).

Pour comprendre la différence, il est utile de se représenter plus généralement σ_t^2 comme la variance conditionnelle de ε_t sachant un certain ensemble d'information $I(t-1)$ correspondant à une filtration croissante par rapport à laquelle les processus considérés sont supposés adaptés, moyennant un léger changement de notation :

$$\sigma_t^2 = f_{t-1} = \mathrm{Var}[\varepsilon_t/I(t-1)] = E[\varepsilon_t^2/I(t-1)].$$

pour signifier que c'est le processus f_t qui est directement adapté à la filtration $I(t)$.

Pour analyser, comme le suggère LM, « l'effet d'un choc sur la variance conditionnelle », le plus naturel est de s'intéresser directement à la dynamique de celle-ci, que l'on peut imaginer représentable par un modèle autorégressif d'ordre un :

$$f_t = \omega^* + \gamma^* f_{t-1} + \eta_t \text{ avec } E[\eta_t/I(t-1)] = 0 \qquad (**)$$

Il en résulte immédiatement que :

$$E[\varepsilon_t^2 - \omega^* - \gamma^* \varepsilon_{t-1}^2/I(t-2)] = 0$$

ce qui implique en particulier que $(\varepsilon_t^2 - \omega^* - \gamma^* \varepsilon_{t-1}^2)$ est un processus moyenne mobile d'ordre un que l'on peut écrire :

$$\varepsilon_t^2 - \omega^* - \gamma^* \varepsilon_{t-1}^2 = \nu_t - \beta^* \nu_{t-1}$$

où ν_t est un bruit blanc au second ordre. Cependant la représentation (**) est plus restrictive au sujet de ν_t puisqu'elle garantit en outre que :

$$E[\nu_t - \beta^* \nu_{t-1}/I(t-2)] = 0,$$

ce qui apparaît comme une propriété intermédiaire entre :

[ν_t différence de martingale] et [ν_t bruit blanc au second ordre].

Autrement dit, on a les relations d'implication suivantes dont les réciproques sont fausses (avec $\gamma = \gamma^*$, $\beta = \beta^*$, $\alpha = \gamma^* - \beta^*$) :

$$\varepsilon_t \text{ GARCH semi fort}$$
$$\Longrightarrow$$
$$f_{t-1} = E[\varepsilon_t^2/I(t-1)] \text{ processus AR(1) au sens de } (**)$$
$$\Longrightarrow$$
$$\varepsilon_t \text{ GARCH faible.}$$

On vérifie de plus que le cas GARCH semi fort correspond à un processus AR(1) (**) très particulier puisqu'il signifie que le processus d'innovation η_1 de cet AR(1) est de la forme :

$$\eta_t = \alpha[\varepsilon_t^2 - f_{t-1}].$$

Autrement dit, dans le cas d'un GARCH semi fort, la variance conditionnelle f_t est une fonction déterministe du passé (ε_s, $s \leqslant t$) du processus observable ε, ce qui n'est pas le cas en général puisque rien ne garantit, dans (**), que l'innovation η_t soit parfaitement corrélée à ε_t^2, conditionnellement à l'information passée $I(t-1)$ qui est en général bivariée :

$$I(t-1) \text{ contient la tribu engendrée par les } (\varepsilon_s, \eta_s),\ s < t.$$

En ce sens, la variance conditionnelle f_t peut être vue comme une chaîne de Markov cachée utilisée pour définir la dynamique des observables ε_t. Ce processus de Markov peut par exemple être défini par une EDS (à tendance linéaire pour être conforme à (**)) :

$$df_t = k(\sigma^2 - f_t)dt + h(f_t)dW_t$$

où W est un mouvement brownien standard et $\gamma = \exp(-k)$. Ainsi la structure AR(1) (**) pour la variance conditionnelle est compatible avec un modèle à volatilité stochastique en temps continu sous jacent. A fortiori, la structure GARCH faible l'est aussi , comme l'avaient remarqué Drost et Werker (1996). En revanche, la structure GARCH usuelle à la Engle/Bollerslev (GARCH semi fort) ne l'est pas en général car cette structure n'est pas robuste à l'agrégation temporelle (Drost et Nijman (1993)) .

Cette conclusion plaide pour l'utilisation des GARCH faibles, ou, pour avoir une structure probabiliste plus riche (en particulier pour que f_t s'interprète bien comme une variance conditionnelle et pas seulement comme une projection linéaire), pour l'utilisation des modèles à **volatilité stochastique**, au sens où la dynamique de la volatilité met en œuvre un aléa qui n'est pas parfaitement corrélé au processus d'intérêt ε.

4. VOLATILITÉ STOCHASTIQUE À MÉMOIRE LONGUE EN TEMPS CONTINU

La théorie moderne de l'évaluation par arbitrage des actifs financiers a popularisé l'utilisation des EDS pour décrire la dynamique des cours ou des taux. On pourrait donc imaginer d'étendre cette approche à l'utilisation des EDS fractionnaires en représentant le cours ou le taux $X(t)$ à la date t par une équation du type :

$$dX(t) = -kX(t)\ dt + \sigma\ dW_d(t)$$

où $W_d(t)$ est le Brownien fractionnaire. CR(96) ont développé cette approche en montrant qu'elle permettait en particulier de généraliser les modèles usuels de la structure par termes des taux d'intérêt pour rendre compte de la variabilité des taux longs, plus importante qu'on ne parvient à l'expliquer avec un modèle usuel : un taux long pouvant être interprété comme une moyenne de taux courts anticipés, la mémoire longue est utile pour expliquer que l'aléa ne soit pas complètement éliminé par cet effet de moyenne. Une explication de ce type avait déjà été proposée à travers un modèle en temps discret par Backus et Zin (1993) (cité par LM). Il y a cependant un inconvénient majeur à utiliser le mouvement Brownien fractionnaire comme processus d'innovations d'un prix ou d'un taux : ce processus n'étant pas une semi martingale, il ne peut pas être équivalent à une martingale; on sait donc par la théorie moderne de l'évaluation par arbitrage qu'un tel processus est incompatible avec l'hypothèse fondamentale d'absence d'opportunité d'arbitrage[1].

Comme il n'y a pas par ailleurs énormément de résultats empiriques démontrant de façon convaincante la présence de mémoire longue dans les cours et les taux (voir cependant la revue de ces résultats proposée par LM), on préférera se concentrer ici sur la mémoire longue dans les processus de volatilité. En effet, outre que cet effet mémoire longue est utile, comme bien expliqué par LM, pour rendre compte de l'évidence empirique selon laquelle la volatilité est certes stationnaire mais revient lentement à son niveau moyen (les chocs sur la volatilité sont très persistants), on peut montrer que cette approche est très utile pour l'évaluation des options longues. Les deux phénomènes sont d'ailleurs liés : c'est précisément parce que la volatilité ne revient que très lentement à son niveau moyen qu'elle apparaît encore stochastique, même dans les options très longues où il y a pourtant un effet moyenne comme dans la structure par termes des taux d'intérêt.

Cette idée est développée dans CR(98) à partir d'un modèle à volatilité stochastique du cours $S(t)$ (d'une action par exemple) sous-jacent à l'option :

$$dS(t)/S(t) = \mu(t, S(t))dt + \sigma(t)\ dW^1(t)$$
$$d(\ln \sigma(t)) = k[\theta - \ln \sigma(t)]dt + \gamma\ dW^2(t)$$

Comme expliqué au point 3. ci-dessus, le principe de la volatilité stochastique consiste à introduire un processus d'innovation dans la dynamique markovienne de la volatilité (ici le brownien $dW^2(t)$) qui ne soit pas parfaitement corrélé avec le processus d'innovation du rendement (ici le brownien $dW^1(t)$). Pour la simplicité, on a considéré un processus de volatilité log-normal (trend linéaire en $\ln \sigma(t)$ plutôt qu'en $\sigma^2(t)$). Dans le cas extrême où les deux browniens sont indépendants, la formule de Hull-White d'évaluation à la date t d'une option européenne d'échéance T s'obtient alors en prenant l'espérance conditionnelle (par rapport à la probabilité risque neutre) à la valeur $\sigma(t)$ de la volatilité instantanée à la date t de la formule de Black et Scholes dans laquelle le terme constant σ^2 (où σ est le paramètre de volatilité constante) est remplacée par la moyenne sur la période du processus stochastique de variance conditionnelle :

$$V_t^T = 1/(T-t) \int_{[t,T)} \sigma^2(u)\ du.$$

Si donc le prix de l'option est conforme à ce modèle et l'on calcule une volatilité implicite de Black et Scholes $\sigma^2_{\text{imp}}(t, T)$ en égalant le prix de l'option à la

1. (NDLR) Voir aussi la contribution de Pierre Bertrand dans cette même discussion.