

Chapitre 2

UNE ESTIMATION DYNAMIQUE DE LA LIQUIDITÉ

INTRODUCTION

La mise en place des systèmes électroniques au milieu des années quatre-vingts a favorisé le développement des marchés financiers. De plus, l'informatisation a permis aussi de créer et diffuser des bases de données de haute fréquence. La première approche privilégiée par les chercheurs a été d'agréger les données sur des intervalles réguliers de durée réduite, ce qui permettait d'établir de nouvelles variables telles que des rentabilités par unité de temps. Cependant aucun crédit n'était réellement porté à la variable temps dans la mesure où la plupart des modèles théoriques était des modèles séquentiels tel que le modèle de Kyle (1985).

La notion de temps n'a commencé à être réellement reconnue que lorsque l'on a mis en relief le fait selon lequel les décisions des intervenants n'étaient pas indépendantes du temps. Ainsi Diamond et Verrechia (1987) ont considéré le temps comme une source d'information. Selon eux, l'absence de transaction signifie qu'il y a une mauvaise nouvelle. En revanche, Easley et O'Hara (1992b) interprètent l'absence de transaction par l'absence d'arrivée d'informations nouvelles sur le marché, ce qui signifie que la variable temps n'est pas indépendante du processus de prix.

Devant la multiplication des informations horodatées (les cotations, la taille des échanges, la fourchette, ...), les méthodes d'estimation classiques sont alors devenues caduques. Celles-ci ont laissé peu à peu la place à l'économétrie des données à haute fréquence¹ et ses modèles non paramétriques traitant des événements irrégulièrement espacés dans le temps. Deux nouvelles approches sont alors apparues : les approches par déformation du temps (appelées aussi processus subordonnés) et les modèles de durées.

L'approche par déformation du temps n'est pourtant pas nouvelle dans la mesure où elle avait été introduite, de manière théorique par Mandelbrot et Taylor (1967) ainsi que par Clark (1973). Ces auteurs considéraient en effet qu'il était nécessaire d'analyser le processus de prix à l'aide d'une échelle de temps déformée ; en effet, si les modèles dynamiques usuels ne semblent pas fonctionner correctement en temps calendaire, cela doit signifier qu'ils sont écrits dans un mauvais référentiel de temps. Il est fait alors

¹ Voir Goodhart et O'Hara (1996) ainsi que Gavridis (1998).

référence aux notions de « *market time* » ou encore de « *Business Time* ». L'idée principale est que l'activité du marché, mesurée par des séries comme le volume de transactions, détermine l'échelle de temps intrinsèque des processus stochastiques de prix ou de rentabilité². Les bases de données étant désormais disponibles, de nombreux auteurs³ ont alors repris et développé la notion de déformation du temps en proposant différentes échelles de temps. Cette approche semble cependant être toutefois difficile d'application dans la mesure où on ne sait pas clairement ce qu'est réellement le temps intrinsèque.

L'autre approche, constituée par les modèles de durée, consiste à utiliser une échelle de temps déformée de manière à récupérer le processus des durées entre évènements de cotation. Il s'agit justement de l'approche que nous allons développer au cours de ce chapitre.

En effet, l'objet de ce chapitre est d'adapter au marché français, la mesure de liquidité, nommée VNET, proposée par Engle et Lange (1997). Cette modélisation nécessite la prise en compte de différentes variables dont la durée entre les évènements de cotation. Une estimation de ces durées n'est rendue possible que par l'utilisation des modèles de durée conditionnelle autorégressive, introduits par Engle et Russell (1994).

Ainsi, ce deuxième chapitre se décompose de la manière suivante : la première section a pour objet une revue de la littérature afférente aux différentes modélisations de la durée ; nous concentrons notre examen sur les modèles de durée conditionnelle autorégressifs dont nous présentons différentes estimations et un cas d'application : la modélisation de la liquidité introduite par Engle et Lange (1997). La seconde section possède une vocation empirique puisqu'elle consiste à mettre en œuvre la mesure dynamique de la liquidité, VNET, proposée par Engle et Lange (1997) après avoir déterminé le modèle de durée qu'il convient d'appliquer au marché français.

² De plus, cette appréciation était confortée par quelques faits : les lois marginales des rentabilités ont de plus faibles queues de distribution en temps de cotation qu'en temps calendaire ; il y a moins de volatilité en temps de cotation qu'en temps calendaire (ce qui est logique dans la mesure où l'on rejette les deux aléas que sont la durée de transaction et la danse de la fourchette) ; la dynamique (linéaire) est plus simple en temps de cotation qu'en temps calendaire.

³ Voir Muller et al. (1994), Dacarogna et al. (1994), Ané et Geman (1996, 2000), Ghysels, Gouriéroux et Jasiak (1995) par exemple.

SECTION I - ANALYSE THÉORIQUE

Dynamiques complexes, non-linéarité, non-stationnarité et données irrégulièrement espacées ont empoisonné la vie de plus d'un chercheur. À l'exception de l'analyse économétrique des processus de transition [Lancaster (1990)], les modèles de séries temporelles assurent que l'intervalle de temps qui sépare les observations est égal à travers le temps, ce qui n'est pas le cas avec les données à haute fréquence.

Afin de fournir une structure économétrique pour l'analyse des données irrégulièrement espacées, Engle et Russell (1994) ont proposé le modèle ACD (*Autoregressive Conditional Duration*, i.e. Durée Conditionnelle Autoregressive) qui combine à la fois fonction de hasard et modèle ARCH. En effet, les auteurs proposent de traiter les données à haute fréquence comme un processus de points marqués⁴. Un tel processus correspond à une série de dates d'arrivée avec des caractéristiques associées à ces états d'arrivée. Depuis la présentation de ce modèle, les analyses empiriques de durée entre événements de cotation se sont rapidement développées dans différentes directions et de plus ont intégré des hypothèses de microstructure.

Le bénéfice essentiel des données à haute fréquence est qu'elles fournissent une base appropriée pour tester empiriquement des théories et des hypothèses de microstructure effectuées alors qu'il n'existait pas de telles données. Un sujet commun des modèles théoriques récents de microstructure est que la durée entre différents événements telles que les transactions contient de l'information qui affecte le comportement des acteurs, et ainsi le processus de formation des prix.

En conséquence, les premières études destinées à utiliser la durée comme variable aléatoire ont été consacrées à l'analyse de l'occurrence des transactions. Outre Engle et Russell (1994), Rydberg et Shephard (1998) analysent les arrivées des ordres à l'aide d'un processus de Cox. Hautsch (1999) examine le temps entre les transactions à l'aide d'un modèle semi-paramétrique de hasard. Gouriéroux, Jasiak et Le Fol (1999) introduisent des durées basées sur des mesures d'activité. Ils définissent différentes classes de durée comme les durées pondérées par les volumes. Ceci permet d'illustrer la dépendance entre les durées, les volumes et les prix.

⁴ Nous remercions Nour Meddahi pour ses explications précises de ces processus.

Conjointement à l'analyse des transactions, de nombreux auteurs⁵ se sont intéressés à la dynamique des prix. On ne se focalise plus sur la durée existante entre deux transactions mais sur celle qui sépare deux mouvements de prix. Cette « durée-prix », comme le démontrent Engle et Russell (1994), peut-être considérée comme une fonction inverse du modèle de volatilité. En effet, une « durée-prix » mesure le temps nécessaire à une variation de cours d'une unité. Plus le temps nécessaire pour obtenir une variation de prix est long, moins le marché est volatil. Cette mesure de volatilité est considérée par de nombreux chercheurs comme un modèle GARCH adapté aux données à haute fréquence. Outre les développements du modèle ACD que nous examinons ultérieurement, d'autres modèles ont vu le jour pour décrire la volatilité en séance. Ghysels et Jasiak (1998) développent un modèle ACD-GARCH bivarié qui prend en compte les interactions entre les volatilités des rentabilités passées et les durées⁶. Bauwens et Veredas (1999) formulent une variante de ce modèle : le SCD (*Stochastic Conditional Duration*) qui est le pendant des modèles à volatilité stochastique⁷. Ghysels, Gouriéroux et Jasiak (1999) déterminent un modèle à volatilité stochastique SVD (*Stochastic Volatility Duration*) qui introduit une dépendance dans les durées.

Ces différents modèles de volatilité vont permettre une meilleure analyse des différents processus intra-journaliers. Parmi les différents intérêts de telles modélisations, on peut citer l'ajustement des couvertures de portefeuille [Prigent, Renault et Scaillet (1999)]. Elles peuvent aussi être très utiles pour tester des hypothèses de microstructure [Engle et Russell (1994), Bauwens et Giot (1999)] ainsi que pour permettre d'évaluer la liquidité comme le proposent Engle et Lange (1998).

Nous présentons d'abord différentes spécifications du modèle ACD, modèle précurseur en matière d'analyse de données à haute fréquence. Ensuite, nous analysons l'une des premières applications du modèle ACD, à savoir la construction de la statistique VNET proposée par Engle et Lange (1997).

⁵ Engle et Russell (1994), Engle (1996), Bauwens et Giot (1998), Ghysels et Jasiak (1998), Ghysels, Gouriéroux et Jasiak (1999), Gramming et Maurer (1998) ou encore Meddahi, Renault et Werker (1998).

⁶ Il peut ainsi être défini comme un modèle GARCH avec des coefficients aléatoires où la durée entre les transactions détermine la dynamique des paramètres.

⁷ Ce modèle a pour fondement l'hypothèse selon laquelle les durées sont générées par un facteur stochastique latent qui engendre un processus autorégressif de premier ordre.

I.1 DONNÉES HAUTE FRÉQUENCE ET MODÈLES DE DURÉE

Le modèle ACD (Durée Conditionnelle Autorégressive) a provoqué un renouveau en matière de recherche en économétrie de la finance. Ce premier modèle appliqué aux données à haute fréquence constitue le socle de nombreuses propositions citées précédemment. Nous allons d'abord examiner ce qu'est réellement le modèle ACD d'un point de vue économétrique, et mettre ainsi en relief son apparente ressemblance avec les modèles GARCH. Ce premier modèle est devenu très rapidement une classe de modèles ACD dans la mesure où son succès a entraîné de nombreuses propositions d'extension. Ce phénomène est d'autant plus amplifié que beaucoup de chercheurs adaptent des spécifications GARCH à la classe de modèles ACD. Nous faisons le point de la situation sur ces extensions de manière relativement succincte dans la mesure où les propositions se renouvellent très rapidement ce qui fait que nous ne pouvons prétendre à l'exhaustivité.

I.1.1 Les modèles de Durée Conditionnelle Autorégressive

Engle et Russell (1994, 1998) proposent une nouvelle classe de modèles dynamiques nommés ACD en considérant les données à haute fréquence comme un processus de points marqués (*marked point process*). Ce processus discret offre la possibilité aux dates d'arrivée d'être considérées comme endogènes si le chercheur désire spécifier leur distribution conditionnelle.

Les hypothèses

La modélisation ACD repose sur deux hypothèses. Tout d'abord, Engle et Russell (1998) supposent que la durée entre les deux événements de cotation peut être décomposée en deux parties : une première composante déterministe et une seconde aléatoire. On suppose que la relation existante entre les deux composantes est multiplicative.

Soit N , le nombre d'événements de prix qui se produit de manière aléatoire durant la séance. Ces N événements sont notés $i = 1, \dots, N$ du premier observé au dernier.

Soit X_i , la durée qui correspond au temps séparant le $i^{\text{ème}}$ et le $(i-1)^{\text{ème}}$ événement.

Nous avons donc $X_i = t_i - t_{i-1}$.

Soit y_i la composante déterministe qui n'est autre que l'espérance conditionnelle de la durée sachant son passé : $y_i = E[X_i / X_{i-1}]$

Soit e_i la composante aléatoire ;

La relation multiplicative peut alors s'écrire de la manière suivante :

$$X_i = y_i * e_i$$

La seconde hypothèse concerne la composante aléatoire de la durée. Engle et Russell (1994) supposent que cette composante est identiquement et indépendamment distribuée.

$$e_i = \frac{x_i}{y_i} \text{ sont } i.i.d$$

Ceci signifie que l'espérance conditionnelle de la durée (la composante déterministe) doit capter la dépendance des durées dans le temps.

Ceci amène les auteurs à suggérer une structure autorégressive dans la spécification de l'espérance conditionnelle. Nous obtenons donc le modèle $ACD(p, q)$ suivant :

$$y_i = w + \sum_{j=1}^p a_j * X_{i-j} + \sum_{k=1}^q b_k * y_{i-k} \quad (2.1)$$

Modélisation de la vraisemblance

L'hypothèse concernant la spécification de la densité conditionnelle de la durée nous permet de pouvoir estimer la log-vraisemblance de la manière suivante :

$$L(X_1, \dots, X_{N(T)}; \mathbf{q}) = \sum_{i=1}^{N(T)} \text{Log } f(X_i / X_1, \dots, X_{i-1}; \mathbf{q}) = y_i$$

$$\text{et } y_i = \int_0^{N(t)} X_i \cdot g(X_i / X_1, \dots, X_{i-1}) d_i$$

où $g(X_i / X_1, \dots, X_{i-1})$ indique la fonction de densité conditionnelle de la durée, tandis que \mathbf{q} est un aléa.

Selon l'hypothèse liée à la composante aléatoire de la durée, les e_i correspondent aux erreurs, mais ils représentent aussi les durées standardisées.

Si nous divisons les durées par la fonction d'espérance conditionnelle de la durée, les résidus obtenus sont théoriquement indépendamment et identiquement distribués.

De cette manière la vraisemblance devient :

$$L(X_1, \dots, X_{N(T)}; \mathbf{q}) = \sum_{i=1}^{N(T)} \text{Log } f(\mathbf{e}_i; \mathbf{q}) = \sum_{i=1}^{N(T)} \text{Log} \left[\mathbf{I}(t/N(t), t_1, \dots, t_{N(t)}) * e^{(-\mathbf{e}_i)} \right]$$

La fonction de Log-vraisemblance va alors dépendre de la spécification de \mathbf{e}_i et bien entendu de la modélisation de l'intensité conditionnelle.

Selon la spécification retenue, il est désormais possible de retrouver la fonction d'intensité conditionnelle des transactions. Pour cela, il est nécessaire d'introduire la notion de fonction de hasard. A l'instant t , elle s'interprète comme étant la probabilité de constater l'existence d'une transaction à l'instant t sachant qu'il n'y a pas eu de transactions depuis une certaine durée. Analytiquement, la fonction de hasard correspond au rapport entre la fonction de densité de la durée et la fonction de survie de cette même durée. La fonction de Survie (*Survivor function*) $S(y)$ caractérise la probabilité de rester en l'état initial durant au moins y unités de temps.

Soit $I_0(t)$, la fonction de hasard de base pour les durées standardisées, $P_0(t)$ la densité correspondante, et $S_0(t)$ la fonction de survie associée, nous obtenons alors :

$$I_0(t) = \frac{P_0(t)}{S_0(t)}$$

Quelles que soient les distributions des durées (standardisées ou non), nous pouvons écrire l'intensité conditionnelle des événements de cotation de la manière suivante :

$$\mathbf{I}(t/N(t), t_1, \dots, t_{N(t)}) = I_0(t) \left\{ \frac{t - t_{N(t)}}{\mathbf{Y}_{N(t)+1}} \right\} \frac{1}{\mathbf{Y}_{N(t)+1}}$$

La forme de la fonction de hasard dépend alors du choix de la loi suivie par les différentes durées. Ceci entraîne la proposition de plusieurs modèles de durée conditionnelle autorégressive : EACD, WACD, BURR-ACD. Nous allons présenter ces différents modèles après avoir analysé le traitement des saisonnalités.

Le traitement de la saisonnalité

La classe de modèles ACD introduite doit être adaptée aux saisonnalités observées sur les marchés financiers. En effet, de nombreux auteurs ont pu constater qu'il existait des saisonnalités aussi bien en termes d'activité⁸, de rentabilité ou encore de volatilité [Jain et Joh (1988), Foster et Viswanathan (1990), Hamon et Jacquillat (1992),...]. Ainsi, l'espérance de la durée diffère selon l'instant de la séance où nous nous situons. Ainsi, en termes d'activité, les durées entre transactions sont plus courtes à l'ouverture et à la clôture de la séance, et sont plus longues durant la pause déjeuner [Gourieroux, Jasiak et Le Fol (1999)].

Afin que les modèles ACD prennent en compte le facteur saisonnier, il est donc nécessaire de décomposer l'espérance de la durée en deux composantes selon une relation multiplicative : une composante déterministe fonction du temps et une composante stochastique. La première étape de la modélisation ACD consiste donc à désaisonnaliser les durées.

Ces durées désaisonnalisées, notées \tilde{X}_i , sont donc estimées de la manière suivante :

$$\tilde{X}_i = \frac{X_i}{f(t_{i-1}; \mathbf{q}_f)}$$

Par conséquent l'espérance conditionnelle de la durée s'écrit de la manière suivante :

$$E(X_i / X_1, \dots, X_{i-1}) = f(t_{i-1}; \mathbf{q}_f) * \mathbf{y}_i(\tilde{X}_1, \dots, \tilde{X}_{i-1}; \mathbf{q}_y)$$

L'hypothèse concernant l'indépendance de la distribution des erreurs (i.e. les durées standardisées) est préservée. Nous obtenons donc :

$$\mathbf{e}_i = \tilde{X}_i / \mathbf{y}_i(\tilde{X}_1, \dots, \tilde{X}_{i-1}; \mathbf{q}_y)$$

Il se pose alors le problème de savoir quelle méthode de désaisonnalisation il faut employer. La méthode initiale proposée par Engle et Russell (1994) consiste à estimer des fonctions *splines* linéaires à partir de moyennes horaires. Un tel choix n'est pas sans présenter plusieurs inconvénients : ils sont au nombre de trois.

⁸ Voir par ailleurs le premier chapitre de cette thèse.

La fonction *spline* linéaire va révéler des déséquilibres au niveau des nœuds. Ces déséquilibres vont directement affecter la fonction *spline*. L'hypothèse d'une fonction *spline* linéaire entre les nœuds est restrictive, particulièrement lorsqu'il y a des changements de régime de type ouverture clôture. Ce problème est d'autant plus important que l'intervalle de temps est grand. Ainsi le risque de constater l'existence de plusieurs points d'inflexion entre les nœuds est élevé.

L'hypothèse selon laquelle le profil du processus est identique chaque jour semble restrictive. Cela est le cas lors des séances spéciales telle que la séance d'expiration des contrats à terme [Le Saout (1999a)].

Les deux premiers problèmes peuvent être réglés en utilisant une fonction *spline* polynomial. Buse et Lim (1977) démontrent⁹ comment estimer une fonction *spline* cubique en utilisant la méthode des moindres carrés restreints. Les restrictions consistent désormais à déterminer l'identité des dérivées première et deuxième de la fonction *spline* au niveau des nœuds afin de s'assurer de la continuité.

Le dernier problème peut être résolu en intégrant des variables *dummies* journalières ou alors en sélectionnant dans un autre échantillon les séances spéciales. Cette dernière solution, en dépit de son extrême simplicité, est certainement la plus pertinente, car elle permet d'obtenir la fonction *spline* appropriée pour chaque séance.

La spécification de base

Étant donné les hypothèses effectuées, le modèle $ACD(p, q)$ général peut s'écrire de la manière suivante :

$$x_i = \mathbf{y}_i * \mathbf{e}_i \quad \text{où } \mathbf{e}_i \text{ est i.i.d.}$$

$$\mathbf{y}_i = \mathbf{w} + \sum_{j=1}^p \mathbf{a}_j * X_{i-j} + \sum_{k=1}^q \mathbf{b}_k * \mathbf{y}_{i-k} \quad (2.1)$$

Les contraintes de signe des coefficients \mathbf{a} et $\hat{\mathbf{a}}$ dépendent des lois de distribution suivies, tandis que le coefficient \mathbf{w} est toujours strictement positif quelles que soient les lois de distribution.

D'un point de vue statistique, nous pouvons constater que les deux premiers moments sont variables dans le temps car ils sont calculés à partir de l'espérance conditionnelle des durées, et ceci quelle que soit la fonction de hasard. Ainsi, par définition, nous avons une espérance conditionnelle égale à y_i . L'espérance non conditionnelle des durées, notée m , peut alors s'écrire selon Engle et Russell (1997) :

$$E(x_i) = m = \frac{w}{1-a-b}$$

Engle et Russel (1997) démontrent ce résultat, et indiquent que la condition nécessaire et suffisante pour valider l'existence de la moyenne est la présence de toutes les racines du polynôme caractéristique associées en dehors du cercle de racine unité¹⁰.

La variance conditionnelle est égale à ψ_i^2 tandis que la variance non conditionnelle s'écrit de la manière suivante :

$$s^2 = m^2 \left(\frac{1-2ab-b^2}{1-(a+b)^2-a^2} \right)$$

Il faut noter que lorsque le coefficient a est positif, l'écart-type non conditionnel est supérieur à la moyenne, ce qui signifie qu'il y a un excès de dispersion¹¹. Il existe en outre un lien étroit entre les deux premiers moments, ce qui conduit Ghysels, Gouriéroux et Jasiak (1998) à proposer le modèle SVD (*Stochastic Volatility Duration*).

Ces résultats préliminaires étant établis, nous allons pouvoir exposer plusieurs modèles de durée conditionnelle autorégressive qui diffèrent notamment par la loi de distribution des différents processus de durée.

I.1.1.1 Le modèle de durée conditionnelle exponentielle

Le premier modèle ACD proposé par la littérature est le modèle EACD, i.e. *Exponential Autoregressive Conditional Duration* [Engle et Russell (1994)]. Dans ce contexte, nous

⁹ Voir aussi Eubank (1988).

¹⁰ Cette hypothèse de stationnarité est validée par la conception même du modèle alors que de sérieux doutes existent en la matière pour ce qui est des données à haute fréquence.

¹¹ Pour une preuve de ce résultat, voir Engle et Russell [1998]

faisons l'hypothèse selon laquelle les durées sont conditionnellement exponentielles c'est-à-dire que les \mathbf{e}_i , i.e. les durations standardisées, suivent une loi exponentielle.

Ainsi les deux équations suivantes caractérisent le modèle EACD :

$$X_i = \mathbf{f}_i * \mathbf{e}_i \text{ avec } \mathbf{e}_i \text{ suit une loi exponentielle}$$

$$\mathbf{y}_i = \mathbf{w} + \sum_{j=1}^p \mathbf{a}_j * \tilde{X}_{i-j} + \sum_{k=1}^q \mathbf{b}_k * \mathbf{y}_{i-k} \quad (2.2)$$

Le modèle EACD(1,1) s'écrit alors :

$$\mathbf{y}_i = \mathbf{w} + \mathbf{a} * \tilde{X}_{i-1} + \mathbf{b} * \mathbf{y}_{i-1} \quad (2.3)$$

Les contraintes de signe sont les suivantes : les coefficients \mathbf{a} et \hat{a} sont positifs ou nuls ; de plus, la somme de ces coefficients doit être inférieure ou égale à l'unité. Comme nous l'avons signalé précédemment, le coefficient \mathbf{w} demeure strictement positif. Ceci permet ainsi de valider les conditions nécessaires et suffisantes en matière de stationnarité telles que le prévoit le corollaire de Lee et Hansen (1994).

Ainsi ce modèle reflète bien le caractère autorégressif des durées. En raison du signe positif ou nul des coefficients \mathbf{a} et \hat{a} , une durée courte, respectivement longue, passée entraîne l'apparition d'une durée courte, respectivement longue, présente. Le phénomène de grappes de transactions est ainsi expliqué par le modèle.

L'espérance conditionnelle de la durée se calcule alors de la manière suivante :

$$E(X_i / I_{i-1}) = \mathbf{f}_i * 1 = \mathbf{f}_i$$

où l'information passée $I_{i-1} = (X_{i-1}, \mathbf{y}_{i-1})$

Or, par définition, l'espérance conditionnelle de la durée est aussi égal à \mathbf{y}_i . Par conséquent, nous avons :

$$\mathbf{f}_i = \mathbf{y}_i \text{ et } \mathbf{e}_i = \begin{pmatrix} X_i \\ \mathbf{y}_i \end{pmatrix}$$

Dans le cas, la fonction de hasard de base associée à la durée standardisée, est égale à l'unité. Dès lors, la fonction d'intensité conditionnelle est estimée ainsi :

$$I(t / X_1, \dots, X_{N(t)}) = \mathbf{y}_{N(t)+1}^{-1}$$

Or, cette fonction de densité correspond à la fonction de hasard de la durée. En conséquence, la fonction de densité de la durée s'exprime de la manière qui suit :

$$f(X_i / X_{i-1}) = I(t / X_1, \dots, X_{i-1}) * e^{(-e_i)} = \frac{e^{-\frac{X_i}{\mathbf{y}_i}}}{\mathbf{y}_i}$$

La fonction Log Vraisemblance peut s'écrire alors :

$$L(X_1, \dots, X_{N(t)}) = \sum_{i=1}^{N(t)} -\ln(\mathbf{y}_i) - \left(\frac{X_i}{\mathbf{y}_i} \right)$$

I.1.1.2 Le Modèle de durée conditionnelle de type weibull

Engle et Russell (1994) proposent également une distribution de Weibull $(1, \mathbf{g})$ pour les durations standardisées, d'où la modélisation du WACD, i.e. le *Weibull Autoregressive Conditional Model*. La plus grande flexibilité qu'offre cette distribution, par rapport à la loi exponentielle, provient du fait que la fonction d'intensité conditionnelle devient une fonction à un seul paramètre ; celui-ci va permettre de déterminer l'évolution croissante ou décroissante de la fonction de hasard.

Ainsi les deux équations suivantes composent le modèle WACD :

$$X_i = \mathbf{f}_i * \mathbf{e}_i \text{ avec } \mathbf{e}_i \text{ suit une loi de Weibull } (1, \mathbf{g})$$

$$\mathbf{y}_i = \mathbf{w} + \sum_{j=1}^p \mathbf{a}_j * \tilde{X}_{i-j} + \sum_{k=1}^q \mathbf{b}_k * \mathbf{y}_{i-k} \quad (2.4)$$

Le modèle WACD(1,1) s'écrit alors :

$$\mathbf{y}_i = \mathbf{w} + \mathbf{a} * \tilde{X}_{i-1} + \mathbf{b} * \mathbf{y}_{i-1} \quad (2.5)$$

Les contraintes de signe sont identiques à celles du modèle EACD. les coefficients \mathbf{a} et $\hat{\mathbf{a}}$ doivent demeurer positifs ou nuls, tandis que la somme de ces coefficients doit être inférieure ou égale à l'unité. De même, le coefficient \mathbf{w} demeure strictement positif.

L'espérance conditionnelle de la durée entre les événements de cotation se calcule, puisque la moyenne d'une loi de Weibull $(1, \mathbf{g})$ est $\Gamma(1 + 1/\mathbf{g})$, de la manière suivante :

$$E(x_i / I_{i-1}) = f_i * \Gamma[1 + 1/g] = y_i$$

On en déduit l'égalité suivante :

$$X_i = \frac{y_i}{\Gamma(1+1/g)} * e_i$$

Les durations standardisées peuvent alors être exprimées comme il vient :

$$e_i = \frac{X_i * \Gamma(1+1/g)}{y_i}$$

Compte tenu de la fonction de hasard de base associée à la durée standardisée, la fonction d'intensité conditionnelle est estimée ainsi :

$$I(t / t_1, \dots, t_{N(t)}, N(t)) = \left(\frac{\Gamma(1+1/g)}{y_{N(t)+1}} \right)^g * (t - t_{N(t)})^{g-1} * g$$

A partir de l'estimation de la durée standardisée, la fonction de densité de la durée séparant deux événements prix va s'exprimer comme suit :

$$f(X_i / I_{i-1}) = I(t / X_1, \dots, X_i) * e^{(-e_i)^g}$$

$$\text{c'est-à-dire : } f(X_i / I_{i-1}) = \frac{g}{X_i} * \left(\frac{\Gamma(1+1/g) * X_i}{y_i} \right)^g * e^{\left(\frac{-\Gamma(1+1/g) * X_i}{y_i} \right)^g}$$

Il nous est donc désormais possible d'estimer la fonction de Log Vraisemblance de la manière suivante :

$$L(X_1, \dots, X_{N(t)}) = \sum_{i=1}^{N(t)} \ln \left(\frac{g}{X_i} \right) + g * \ln \left(\frac{\Gamma(1+1/g) * X_i}{y_i} \right) - \left(\frac{\Gamma(1+1/g) * X_i}{y_i} \right)^g$$

Lorsque $g < 1$, la fonction de hasard est décroissante, ce qui signifie que la probabilité d'avoir une durée longue est faible ; inversement, si $g > 1$, alors la probabilité d'une durée plus allongée entre les événements de cotation devient plus élevée. Lorsque $g = 1$, nous nous retrouvons dans le cas particulier où la loi de Weibull correspond à une loi exponentielle (i.e. la fonction de hasard est égale à l'unité).

I.1.1.3 Les modèles de durée conditionnelle de type Burr-ACD

Gramming et Maurer (1998) développent un modèle de la classe des ACD qui est fondé sur une distribution de Burr. Ils estiment que leur modèle offre une plus grande flexibilité au profil de la fonction de hasard conditionnel que les modèles traditionnels EACD et WACD. En outre, la distribution de Burr contient comme cas particulier les distributions exponentielles, Weibull ainsi que la distribution log-logistique dans la mesure où la distribution de Burr correspond à un mélange de distributions Gamma et Weibull¹².

Le modèle général Burr-ACD(p, q) est le suivant :

$$x_i = y_i * e_i \quad \text{où } e_i \text{ suit une loi de Burr.}$$

$$y_i = w + \sum_{j=1}^p a_j * X_{i-j} + \sum_{k=1}^q b_k * y_{i-k} \quad (2.6)$$

Une variable T distribuée selon une loi Burr présente la fonction de densité suivante :

$$f(t) = \frac{m \cdot c \cdot t^{c-1}}{(1 + s^2 \cdot m \cdot t^c)^{h+1}}$$

où $h = s^{-2}$.

La fonction de survie d'une distribution Burr est donnée par :

$$S(t) = (1 + s^2 \cdot m \cdot t^c)^{-h}$$

L'espérance d'une variable distribuée selon une loi Burr est la suivante :

$$E(T) = m^{-1/c} \cdot \frac{\Gamma(1+1/c) \cdot \Gamma(h-1/c)}{s^{2(1+1/c)} \cdot \Gamma(h+1)}$$

L'espérance existe si les paramètres respectent l'inégalité suivante :

$$-c < 1 < cs^{-2}$$

Si s^2 est proche de zéro, alors la distribution de Burr équivaut à une distribution de Weibull et le paramètre c est équivalent au paramètre g de Weibull. Ainsi, pour s^2 et

¹² Voir Lancaster (1990).

c respectivement égaux à zéro et à un, nous allons obtenir une distribution exponentielle. En revanche, si s^2 est égal à un, nous obtenons une distribution Log-logistique.

Gramming et Maurer (1998) définissent alors la transformée suivante de l'espérance conditionnelle de la durée :

$$e_i = y_i \cdot \frac{s^{2(1+1/c)} \cdot \Gamma(h+1)}{\Gamma(1+1/c) \cdot \Gamma(h-1/c)}$$

Avec l'espérance conditionnelle de la durée définie de la manière suivante :

$$y_i = E(\tilde{X}_i / \tilde{X}_{i-1})$$

Les auteurs effectuent alors l'hypothèse selon laquelle les variables aléatoires $\tilde{X}_i \cdot e_i^{-1}$ sont i.i.d. par une loi de Burr avec un paramètre m égal à l'unité.

En ce qui concerne la fonction de Log Vraisemblance, les auteurs proposent la formulation suivante¹³ :

$$L = \sum_{i=1}^{N(t)} \ln(c) - c \ln(e_i) + (c-1) \ln(\tilde{X}_i) - (n+1) \ln(1 + s^2 \cdot e_i^{-c} \cdot \tilde{X}_i^c) [X_i \cdot \Gamma(1+1/g)]$$

Gramming et Maurer (1998) comparent les formulations EACD et WACD avec leur modèle Burr-ACD dans le cadre de l'introduction d'une action¹⁴ sur le système allemand IBIS. Il leur apparaît que ce nouveau modèle donnerait des résultats plus satisfaisants. Ceci n'est guère surprenant dans la mesure où il contient comme cas particulier les distributions exponentielle et de Weibull ; il suffit que les variables traitées ne suivent pas exactement ces lois pour que l'estimation s'avère meilleure. L'intérêt d'un tel modèle réside dans le fait que les hypothèses liées aux fonction de hasard et de survie sont moins restrictives. Cependant, comme nous avons pu le constater, la mise en œuvre n'est guère aisée.

I.1.1.4 Les Modèles Logarithmiques de durée conditionnelle

Bauwens et Giot (1997) proposent une autre modélisation ACD nommé Log-ACD. Contrairement à ce qu'indique son nom, les auteurs maintiennent une distribution de Weibull pour les durations standardisées. La différence intervient dans la relation

¹³ Voir Gramming et Maurer (1998) pour la démonstration.

multiplicative qui lie la composante déterministe à la composante aléatoire. Les auteurs considèrent la relation entre la composante aléatoire et l'exponentielle de la composante qui est déterministe (i.e. la fonction exponentielle de l'espérance de durée conditionnelle). Cette proposition va permettre d'assouplir les contraintes, restrictives jusque là, exercées sur les coefficients \mathbf{a} et $\hat{\mathbf{a}}$, sans pour autant complexifier le modèle par des fonctions de hasard et de survie difficiles à manier puisque la distribution de Weibull est conservée.

De la même manière que pour les modèles ACD de base introduits par Engle et Russell (1998), nous émettons toujours l'hypothèse que la fonction déterministe du temps peut être identifiée, et que celle-ci est une fonction multiplicative. En revanche, nous supposons que c'est l'exponentielle de cette fonction qui intervient de manière multiplicative dans la durée.

Nous obtenons donc la première équation du modèle comme suit :

$$X_i = e^{f_i} * \mathbf{e}_i$$

et, nous savons par hypothèse, que :

$$\mathbf{e}_i = \frac{X_i}{e^{f_i}} \text{ i.i.d}$$

Étant donné que la fonction f_i est proportionnelle au logarithme de l'espérance conditionnelle de la durée, nous pouvons écrire :

$$y_i = \text{Ln}[E(X_i / I_{i-1})]$$

d'où :

$$e^{y_i} = E(X_i / I_{i-1})$$

Si nous effectuons l'hypothèse selon laquelle les durées standardisées suivent une loi de Weibull $(1, \mathbf{g})$, il est possible d'établir une relation entre le logarithme de l'espérance conditionnelle et le logarithme de la fonction déterministe.

$$e^{y_i} = e^{f_i} * \Gamma(1 + 1/\mathbf{g})$$

¹⁴ Deutch Telekom en l'occurrence.

Le logarithme de l'espérance conditionnelle s'écrit donc :

$$\Psi_i = \mathbf{w} + \mathbf{a}g(X_{i-1}, \mathbf{e}_{i-1}) + \mathbf{b}\Psi_{i-1}$$

La seule contrainte qui pèse sur le signe des coefficients concerne \hat{a} . Celui-ci doit être strictement inférieur à l'unité.

Enfin, connaissant les durations standardisées, nous sommes en mesure de déterminer la fonction de densité de la durée :

$$f(X_i) = \frac{\tilde{a}}{X_i} \left(\frac{X_i \cdot \tilde{A}(1+1/\tilde{a})}{e^{\theta_i}} \right)^* e^{\left[-\left(\frac{X_i \cdot \tilde{A}(1+1/\tilde{a})}{e^{\theta_i}} \right)^{\tilde{a}} \right]}$$

Et la Log Vraisemblance va pouvoir être formulée de la manière qui suit :

$$L = \sum_{i=1}^{N(t)} \ln(\mathbf{g}) - \ln(X_i) + \mathbf{g} \ln[X_i \cdot \Gamma(1+1/\mathbf{g})] - \mathbf{g}\Psi_i - \left[\frac{X_i \cdot \Gamma(1+1/\mathbf{g})}{e^{\Psi_i}} \right]^{\mathbf{g}}$$

On peut s'apercevoir que la fonction de Log Vraisemblance va dépendre de la fonction $g(-)$. Bauwens et Giot (1997) proposent deux spécification possibles pour cette fonction.

➤ La première proposition est de considérer la fonction $g(X_{i-1}, \mathbf{e}_{i-1})$ comme étant égale à la fonction $\ln(X_{i-1})$. On obtient ainsi le modèle général Log-ACD₁ :

$$\Psi_i = \mathbf{w} + \sum_{j=1}^p \mathbf{a}_j \ln(X_{i-j}) + \sum_{k=1}^q \mathbf{b}_k * \Psi_{i-k} \quad (2.7)$$

Si on souhaite écrire le modèle en fonction des durées standardisées, il vient :

$$\Psi_i = \mathbf{w}' + \sum_{j=1}^p \mathbf{a}_j \ln(\mathbf{e}_{i-j}) + \sum_{k=1}^q (\mathbf{a}_k + \mathbf{b}_k) * \Psi_{i-k} \quad (2.8)$$

Le modèle Log-ACD₁ (1,1) s'écrit alors :

$$\Psi_i = \mathbf{w}' + \mathbf{a} \ln(\mathbf{e}_{i-1}) + (\mathbf{a} + \mathbf{b}) \Psi_{i-1} \quad (2.9)$$

où $\mathbf{w}' = \mathbf{w} + \mathbf{a} \ln[\Gamma(1+1/\mathbf{g})]$

Ce modèle est le pendant du modèle Log-GARCH introduit par Geweke (1986). Pour respecter l'hypothèse de la stationnarité de la covariance des $Ln(X_i)$, $|\mathbf{a} + \mathbf{b}|$ doit être inférieur à l'unité.

➤ La seconde proposition revient à considérer que la fonction $g(X_{i-1}, \mathbf{e}_{i-1})$ s'apparente à la durée standardisée \mathbf{e}_{i-1} . Le modèle nommé Log-ACD₂ général peut être ainsi formulé :

$$\Psi_i = \mathbf{w} + \sum_{j=1}^p \mathbf{a}_j (\mathbf{e}_{i-j}) + \sum_{k=1}^q \mathbf{b}_k * \Psi_{i-k} \quad (2.10)$$

Le modèle Log-ACD₂ (1,1) s'écrit alors :

$$\Psi_i = \mathbf{w} + \mathbf{a} X_{i-1} \cdot \frac{\Gamma(1+1/g)}{e^{\Psi_{i-1}}} + \mathbf{b} \Psi_{i-1} \quad (2.11)$$

Ou plus simplement :

$$\Psi_i = \mathbf{w} + \mathbf{a} \mathbf{e}_{i-1} + \mathbf{b} \Psi_{i-1} \quad (2.12)$$

Cette proposition apparaît plus adaptée pour les faibles durées : l'espérance conditionnelle dépend de son passé mais aussi de l'erreur de prévision passée. Cette modélisation est fortement inspirée du modèle EGARCH (*Exponential GARCH*) proposé par Nelson (1991). Pour respecter l'hypothèse de la stationnarité de la covariance des Ψ_i , $|\mathbf{b}|$ doit être inférieur à l'unité.

Bauwens et Giot (1999) effectuent des comparaisons entre les différents modèles de durée conditionnelle autorégressive existants. Ils concluent que les différences entre ces modèles ne sont pas significatives statistiquement. Ils s'aperçoivent toutefois que les modèles ACD de base semblent surestimer les longues durées, et sous-estimer les durées courtes ce qui s'explique par la présence du logarithme dans leurs nouvelles formulations.

Nous avons présenté plusieurs modèles de la classe ACD. Ces différents modèles, à l'exception du Log-ACD, se distinguent essentiellement par la distribution des « durées prix » standardisées. Ainsi, nous avons exposé les cas où les lois suivies par les « durées prix » standardisées étaient respectivement des lois exponentielle, de weibull et de Burr.

Bien entendu, il est possible de modéliser¹⁵ ces « durées prix » par d'autres lois telles que la loi gamma, la loi log-logistique ou encore la loi log-normale.

I.1.2 Les extensions aux modèles de durée conditionnelle autorégressive

Les modèles ACD d'origine [Engle et Russell (1994)] ont donné lieu à de nombreuses extensions. Différentes voies se sont ouvertes. Le fait que les modèles ACD soient assimilés à des modèles GARCH adaptés aux données à haute fréquence ont facilité l'avancement des recherches. Ainsi, certains auteurs ont souhaité améliorer l'explication de l'espérance de la durée en intégrant des variables additionnelles, ce qui leur permet par la même occasion de tester certaines hypothèses de microstructure suggérées par la littérature théorique. D'autres chercheurs, se sont orientés dans l'adaptation des extensions des modèles ARCH.

Dans un premier temps, nous allons présenter les variables additionnelles introduites dans la modélisation ACD et dans un second temps, nous allons décrire de manière succincte et non exhaustive quelques nouveaux modèles de la classe ACD.

I.1.2.1 Les modèles de durée conditionnelle et la vérification des hypothèses de microstructure

A l'instar de ce qui a pu être fait avec les modèles GARCH [Lamoureux et Lastrapes (1990), Najand et Jung (1991), Foster (1995)...], les auteurs s'intéressant aux modèles ACD ont souhaité intégrer des variables exogènes dans le processus afin non seulement d'améliorer la prévision mais aussi d'expliquer le caractère autorégressif de la durée entre évènements de cotation. Nous présentons ces différentes variables, les hypothèses de microstructure qui leur sont associées et les résultats obtenus par les chercheurs qui se sont intéressés à cette formulation du problème. Il convient de noter que tous les modèles présentés en premier lieu peuvent intégrer des variables additionnelles, cependant les modèles EACD et WACD convergent difficilement dans la mesure où ces modèles sont contraints par le signe des coefficients.

¹⁵ Voir Le Fol (1998) pour une description complète.

I.1.2.1.1 Le taux de transaction

Cette variable représente le nombre de transactions par seconde. Introduire cette variable exogène dans le processus revient à vérifier si le nombre de transactions qui s'effectue à l'intérieur d'une durée τ influence cette durée (présente ou suivante). Avant d'évoquer l'incidence de l'intensité des échanges sur la volatilité, il convient de s'interroger sur un paradoxe. En effet, si nous jugeons une action illiquide en raison de son faible nombre de transactions, nous pouvons penser que sa volatilité est plus élevée dans la mesure où chaque transaction possède une influence prépondérante sur le niveau de prix. Ceci dit, étant donné qu'il y a peu d'activité, les événements de prix devraient être rares et par conséquent les durées vont être longues. Inversement, les actions liquides font l'objet de beaucoup de transactions par conséquent, l'occurrence de variations de prix « significatives » intervient de manière fréquente. Les « durées - prix » sont alors plus courtes.

Dans la littérature qui s'étoffe depuis quelques années, les variables d'activités sont utilisées au titre de variable *proxy* de l'arrivée d'information sur le marché. Ainsi, Lamoureux et Lastrapes (1990) ou encore Najand et Jung (1991) introduisent le niveau d'activité comme variable explicative au sein de l'équation de la variance conditionnelle dans le cadre d'un modèle GARCH. Les auteurs veulent ainsi justifier le fait que l'information véhiculée par les volumes de transactions influence le niveau de la volatilité. Foster (1995) pour sa part démontre, à l'aide de la méthode des moments généralisés, que l'activité et la volatilité sembleraient être plutôt dirigées par une même variable : le taux de diffusion de l'information. En fait, cette réflexion rejoint quelque peu l'analyse de Blume, Easley et O'Hara (1994), selon qui le volume fournirait davantage d'information sur la qualité des signaux des intervenants du marché.

Le modèle d'Easley et O'Hara (1992) prédit que le nombre de transactions influence le processus de découverte de prix à travers l'information qui se détecte lorsque les transactions se concentrent, ainsi les fourchettes devraient être très fréquemment révisées ainsi que le prix coté. En revanche le modèle d'Admati et Pfleiderer (1988) prédit que le nombre de transactions n'a aucune incidence sur l'intensité des prix (i.e. la probabilité d'un mouvement de prix).

Les résultats obtenus par Engle et Russell (1994) tendent à confirmer les hypothèses émises par Easley et O'Hara (1992) puisque la relation existante entre les « durées - prix » et le taux de transaction passé est significativement négative. Ceci veut dire que la durée entre deux évènements prix est plus courte, i.e. la volatilité est plus élevée, après les périodes de fort taux de transactions. Bauwens et Giot (1999) confirment cette relation négative en analysant l'affichage des fourchettes sur le NYSE. Ces résultats étaient attendus au vu du simple raisonnement effectué en début de présentation.

I.1.2.1.2 La fourchette

Engle et Russell (1994), Engle et Lange (1997) et Bauwens et Giot (1999) proposent d'intégrer la fourchette nominale¹⁶, i.e. l'écart de prix entre le meilleur cours acheteur et le meilleur cours vendeur. Certes la danse de la fourchette n'a aucune incidence dans le processus de modélisation des évènements prix dans la mesure où les différents auteurs retiennent le milieu de la fourchette. Cependant, une large fourchette signifie que l'actif financier présente une volatilité plus élevée : il existe un degré d'incertitude plus important sur la valeur fondamentale, ou valeur d'équilibre, de l'action. Par conséquent, les cotations devraient connaître une variabilité plus élevée. Ainsi, une fourchette large devrait engendrer un plus grand nombre de variations de prix significatives, donc ces évènements de cotations devraient intervenir de manière plus fréquente.

Cette approche corrobore d'ailleurs les résultats d'Engle et Russell (1994), de Bauwens et Giot (1999) ainsi que ceux d'Engle et Lange (1997). Ces auteurs, qui modélisent respectivement un EACD(2,2) et un WACD(1,1), obtiennent des résultats très significatifs ; la largeur de la fourchette est corrélée négativement avec l'espérance de la durée : ceci signifie qu'une action ayant une fourchette de prix étroite (large) présente de longues (faibles) durées entre variations de prix. En revanche, d'après Engle et Russell (1994), la présence de la variable fourchette ne semblerait pas influencer la valeur des coefficients α et $\hat{\alpha}$ mais la fourchette semblerait liée au taux de transaction.

¹⁶ Engle et Russell (1994) et Bauwens et Giot (1999) considèrent la moyenne des fourchettes affichées durant la « durée prix » précédente, tandis qu'Engle et Lange (1997) retiennent la dernière fourchette cotée lorsque le dernier évènement prix intervient.

I.1.2.1.3 La taille moyenne des ordres

Engle et Russell (1994) introduisent aussi le nombre d'actions échangées par transaction qui apparaît lui aussi significatif dans la régression, cependant, il convient de signaler que les auteurs recourent à cette nouvelle variable dans un modèle où deux variables exogènes, la fourchette et le taux de transaction, sont déjà présentes. A notre avis, ceci peut biaiser les résultats dans la mesure où ces trois variables possèdent très certainement un coefficient de corrélation élevé. L'interprétation de la taille moyenne des ordres exécutés apparaît dès lors délicate. Easley et O'Hara (1992) considèrent que sur le marché, il existe deux types d'investisseurs : les initiés et les non initiés. Les agents non informés, appelés aussi « *liquidity traders* », achètent et vendent pour des motifs de liquidité. Ils considèrent que le volume observé peut être décomposé en deux types de volume. Il y a un « volume normal » qui correspond au niveau de liquidité moyen de l'actif financier : il est le fruit des échanges effectués sur l'initiative des agents non informés. La différence¹⁷ entre le volume observé et le « volume normal » détermine le « volume inattendu » qui correspond aux transactions effectuées par les investisseurs informés. Ainsi, la probabilité de négocier avec des intervenants initiés s'accroît lorsque le volume anormal augmente, ce qui entraîne une modification de la fourchette par le teneur de marché (aveugle devant l'investisseur) face à ce problème de sélection adverse.

Bauwens et Giot (1997, 1999) analysent sur le NASDAQ et le NYSE le lien avec le volume anormal et n'obtiennent pas de résultats significatifs sur le NASDAQ. Une des explications avancée par les auteurs est la non prise en compte des saisonnalités quotidiennes tel que l'effet Lundi [Jain et Joh (1988), Foster et Viswanathan (1990)]. Les résultats obtenus sur le NYSE corroborent en revanche l'analyse d'Easley et O'Hara (1992).

En fait, la relation négative entre les durées et la taille de la transaction s'explique aussi sans passer par la théorie des relations prix volume¹⁸. Une transaction de grande taille engendre une baisse de la profondeur du carnet d'ordres, aussi la probabilité de toucher une autre limite de prix s'accroît. Il est à noter que certains auteurs estiment que la taille des ordres n'est pas significative dans la formation du processus de découverte de prix ; ceci s'explique par le fait que les investisseurs informés préfèrent avancer sur le marché à

¹⁷ Il n'est pas ici tenu compte des volumes échangés en raison de la diffusion de l'information publique.

visage couvert en exécutant de petits ordres afin de profiter le plus longtemps possible de leur information privilégiée. C'est un des fondements de l'analyse de Jones, Kaul et Lipson (1994) qui démontrent que c'est l'occurrence des transactions qui joue un rôle sur la volatilité et non pas le volume lui-même endogène au nombre de transactions. Un tel raisonnement constitue peut-être une explication aux résultats de Bauwens et Giot (1997).

I.1.2.1.4 Le déséquilibre

Bauwens et Giot (1997) analysent le processus d'affichage de la fourchette sur le NASDAQ, marché gouverné par les prix. Glosten et Milgrom (1985) construisent un modèle d'information dans lequel les investisseurs et les teneurs de marché ne possèdent pas la même information sur la valeur fondamentale de l'action qu'ils échangent. S'il n'y avait que des investisseurs non informés sur le marché, la fourchette serait réduite et les volumes de transactions « acheteuses » devraient être équivalents aux volumes de transactions « vendeuses ». La situation diffère en présence d'investisseurs initiés ; ces derniers vont acheter ou vendre en fonction de leurs informations : nous devrions alors assister à un déséquilibre entre les ordres acheteurs et vendeurs révélant ainsi la présence d'initiés. Devant ce constat, le *market maker* révisé la fourchette. Selon Bauwens et Giot (1997), la modélisation de la durée entre affichage de fourchette peut expliquer les variations de liquidité durant la séance puisque ce modèle prend en compte le comportement passé du *market maker* (partie autorégressive du modèle) et du déséquilibre actuel entre les achats et les ventes.

Sur le plan empirique, Bauwens et Giot (1997) introduisent cette variable dans le modèle Log-ACD. Leurs résultats indiquent une relation négative très significative. Ceci tend à prouver que les teneurs de marché révisent leurs propositions de prix compte tenu de l'état passé du marché.

I.1.2.2 Une véritable classe de modèles ACD

On va présenter trois modèles de la classe ACD qui sont destinés à améliorer la prévision du temps séparant différents événements de cotation. On met particulièrement en avant la

¹⁸ Voir Karpoff (1997).

présence de plusieurs régimes dans le processus de cotation et la mémoire longue du processus. On conclut par des améliorations possibles de cette classe de modèles ACD.

I.1.2.2.1 Les modèles de durée conditionnelle à seuil

Par analogie au modèle TARARCH (*Threshold ARCH*) [Zakoian (1994), Gouriéroux (1992)], où la variance conditionnelle est une fonction par morceau, susceptible de différentes réactions selon le signe et la grandeur des chocs, Zhang, Russell et Tsay (1999) tentent de modéliser un *Threshold ACD*¹⁹. Pour cela, les auteurs définissent deux états - Une période de forte activité et une période de faible activité - dont laissent déjà augurer les articles d'Engle et Russell (1994) et Hasbrouck (1999) puisque les premiers auteurs cités estimaient les coefficients des variables explicatives supplémentaires (i.e. la fourchette et les volumes) dans un modèle ACD selon le niveau d'intensité des transactions.

Le modèle TACD permet d'identifier la relation non linéaire existante entre l'espérance conditionnelle de la durée entre événements de cotation et les durées passées, et réduit aussi significativement les statistiques de test non linéaire proposées par Engle et Russell (1998).

Il permet aussi de corroborer les hypothèses faites par différents modèles théoriques tel que celui proposé par Easley et O'Hara (1992). Le comportement des variables économiques leur permet d'identifier le régime de forte intensité des échanges comme un régime de transactions initiées par des investisseurs informés ; inversement, le bas régime d'activité est associé à des échanges d'intervenants non informés.

L'hypothèse selon laquelle les erreurs suivent une distribution gamma fournit une fonction de hasard flexible pour chaque régime et résout un problème d'excès de dispersion provoqué par la présence d'hétérogénéité.

Le modèle TACD permet ainsi d'améliorer la compréhension du phénomène d'autocorrélation des durées entre événements de cotation et constitue un pas important vers la modélisation du processus de cotation dans sa totalité. La connaissance des différentes dynamiques (transactions, prix, taux d'intérêt,...) peut ainsi faciliter l'estimation et l'identification de différentes séries fondamentales [Duffie et Glynn (1996)].

I.1.2.2.2 Modèles de durée et persistance

Jasiak (1999) remarque que l'autocorrélation entre les événements de cotation s'apparente à un processus de mémoire longue. Un choc dans le processus de transaction possède un impact qui persiste dans le temps. C'est pourquoi, Jasiak (1999) développe le modèle ACD afin d'en proposer une nouvelle version nommée FIACD (*Fractionally Integrated ACD*) par analogie au modèle IGARCH introduit par Engle et Bollerslev (1986) et au modèle FIGARCH proposé par Baillie, Bollerslev et Mikkelsen (1996).

Engle et Bollerslev (1986) ont effectué la proposition du modèle IGARCH afin d'exprimer l'idée que, dans certains cas, les modèles GARCH puissent se caractériser par une persistance dans la variance, ou encore que les prévisions à plusieurs étapes de la variance conditionnelle, ne convergent pas vers la variance non conditionnelle. Ainsi, les modèles IGARCH font partie d'une classe de modèles « caractérisés par une propriété de variance persistante dans lesquels l'information courante reste importante pour les prévisions des variances conditionnelles à tous les horizons » [Engle et Bollerslev (1986)].

Jasiak (1999) étudie la persistance des chocs dans le processus FIACD dans le cadre d'une analyse de la réaction du marché face à l'introduction d'une innovation²⁰. La fonction réponse²¹ mesure l'effet d'un choc sur le comportement temporel des séries. Le modèle FIACD apparaît significatif ce qui laisse supposer que le modèle de base ACD comporte quelques imperfections. De plus, dans une structure à régime multiple, l'auteur met en relief un lien existant entre la persistance des durées et la dynamique des rentabilités. Ces premiers résultats laissent entrevoir des voies possibles de recherches en matière d'efficience.

I.1.1.3.1 Le modèle Log-ACD asymétrique

Bauwens et Giot (1998) proposent une extension intéressante du modèle Log-ACD qui diffère dans sa forme comparée aux deux extensions présentées précédemment. Les auteurs

¹⁹ Il faut noter que Hamelink (1998) suggère aussi une telle modélisation.

²⁰ Voir aussi Gouriéroux et Jasiak (1999) qui introduisent un concept d'innovation adapté à l'analyse des dynamiques non linéaires ; les auteurs expliquent comment le processus initial peut être exprimé en fonction des valeurs présentes et passées de l'innovation et définissent des fonctions réponses à des chocs transitoires ou permanents.

²¹ L'auteur reprend l'approche de Gouriéroux et Jasiak (1999) à savoir que la fonction réponse a pour fondement une décomposition à la Volterra.

développent en effet un modèle Log-ACD asymétrique. La modélisation de la durée entre deux évènements prix dépend du sens de la variation des prix. On retrouve ainsi l'idée de seuil dans la mesure où les auteurs vont considérer deux états du monde : une variation de prix négative et une variation de prix positive durant la « durée prix ». L'objet de la recherche consiste à découvrir si l'espérance conditionnelle dépend du sens de la variation de prix. Le comportement des investisseurs est-il identique quelle que soit l'évolution du marché ; Diamond et Verrechia (1987) considèrent qu'une « mauvaise nouvelle » se traduit par l'absence de transaction alors qu'une augmentation de la fréquence des échanges signifie une « bonne nouvelle ».

Bauwens et Giot (1998) estiment tout d'abord un modèle de transition appliqué au sens de la variation de prix, ce qui leur permet d'affirmer que la probabilité d'un renversement de tendance est plus faible qu'une continuité de la tendance (i.e. une rentabilité positive au temps t a plus de chance d'être suivie par une rentabilité positive que négative et réciproquement). Les auteurs analysent ensuite l'espérance conditionnelle de la durée. Si on observe une augmentation du prix à l'instant t , la durée d'attente est plus faible si dans l'état antérieur, on avait déjà enregistré une variation positive plutôt qu'une variation négative. En revanche, si dans un instant t , nous avons constaté une baisse de prix, la durée d'attente est là encore, plus faible si dans l'état antérieur, nous avons observé une variation positive. Ces résultats signifient qu'un événement prix positif engendre une nouvelle « durée prix » plus rapidement que si l'événement prix était négatif.

Les résultats de Bauwens et Giot (1997) indiquent clairement qu'il existe des dynamiques différentes selon que le marché évolue à la hausse ou à la baisse. Une telle spécification, tel que le suggère aussi Russell (1999), permet non seulement d'améliorer la prévision des durées mais surtout laisse supposer une utilisation en termes de gestion dynamique de portefeuille, car il ne s'agit pas ici d'estimer seulement le temps entre échanges mais de tenir compte de la rentabilité du marché.

Les extensions à la nouvelle classe de modèles ACD (Durée Conditionnelle Autorégressive) introduite par Engle et Russell (1994) ne s'arrêtent pas là. De nombreuses autres propositions ont ainsi vu le jour depuis cette modélisation. Ainsi des modèles multivariés commencent à être développés [Engle et Lunde (1997), Engle et Russell (1999), Russell (1998, 1999)] afin de prendre en compte conjointement plusieurs processus

d'arrivées d'évènements de transactions tels que les processus de transactions et de modifications de fourchettes ou encore de prix. Les principales limites inhérentes à cette classe de modèle sont du ressort de la statistique [incertitudes sur la stationnarité, corrélation entre les deux premiers moments...], néanmoins, les nouveaux modèles ayant pour fondement la classe ACD, tendent à prendre en compte ces remarques.

Depuis peu, certains auteurs tentent de servir de « l'outil » ACD dans diverses formalisations. Ainsi, Engle et Lange (1997) proposent une nouvelle modélisation de la liquidité en séance à partir de l'analyse de la profondeur. Nous allons présenter leur modèle.

I.2 VNET : UNE MESURE D'ESTIMATION DE LA LIQUIDITÉ

Engle et Lange (1997) proposent une nouvelle mesure de la liquidité intitulée VNET directement liée à la profondeur du marché. Cette mesure est construite à partir des volumes d'achat et de vente en excès durant un événement de marché défini par un mouvement de prix. Étant donné que cette mesure varie dans le temps, on peut tenter de la prévoir et de l'expliquer. Ainsi, leur modélisation consiste à combiner un modèle de durée conditionnelle autorégressive avec une régression incluant des variables explicatives de microstructure afin de mieux expliquer les variations de liquidité du marché durant la séance. Compte tenu des résultats qu'ils obtiennent, ils affirment que l'évolution de la profondeur réelle du marché, contrairement à certains modèles théoriques relatifs aux coûts d'inventaire, dépend avant tout de l'environnement interne du marché.

Nous allons présenter cette nouvelle mesure : tout d'abord, on va décrire les fondations de la mesure à savoir la justification du choix de la profondeur ; dans un deuxième temps, on expose la formalisation de la mesure ; par la suite, on présente les résultats obtenus par Engle et Lange (1997). A partir des travaux de Lange (1998), on analyse la robustesse de la statistique VNET avant de présenter l'influence de la détermination de l'événement prix sur la VNET. Enfin, on conclut par la présentation d'un cas d'école proposé par Lange (1998).

I.2.1 Les fondements de la mesure VNET

Lors des krachs boursiers, la volatilité quotidienne atteint des niveaux très élevés, cependant il n'est pas si évident d'affirmer que le marché est illiquide. En effet, ces journées sont marquées par une très forte activité. Si les vendeurs sont à l'origine d'une grande partie du volume, alors les acheteurs, les fournisseurs de liquidité en pareille situation, absorbent par leurs ordres apparemment un nombre très important d'actions, et on ne peut dès lors être qu'impressionné par la liquidité du marché. Clairement, ni les volumes, ni la volatilité n'apparaissent comme une mesure directe de la liquidité bien qu'elles soient étroitement liées à ce concept.

Engle et Lange (1997) décident donc d'introduire une mesure statistique de la profondeur du marché (VNET). Cette profondeur mesure « combien d'actions peuvent être négociées sans causer de détérioration de prix importante ». Cette nouvelle statistique repère, tout le temps, la dynamique de la liquidité du marché, ce qui a des implications pour la stabilité des marchés et les stratégies de placement des ordres. A titre d'exemple, des gestionnaires de fonds indiciels peuvent tirer partie de cette mesure pour ajuster leur portefeuille, et il en est de même pour les investisseurs qui souhaitent minimiser leurs coûts de transaction.

La liquidité d'un marché est une notion si complexe qu'elle possède une variété de définitions et d'interprétations. Plus simplement, il s'agit de la capacité d'assurer sans coût la vente ou l'achat d'une position. La divergence entre les prix acheteur et vendeur attribuée à la fourchette de prix est la plus connue des facettes de la liquidité. Quoiqu'il en soit, cette mesure indique seulement l'étroitesse du marché pour de faibles volumes. Les ordres de taille importante peuvent être confrontés à un mauvais scénario d'exécution. La déviation de prix peut être très importante pour des volumes élevés et des investisseurs impatientes.

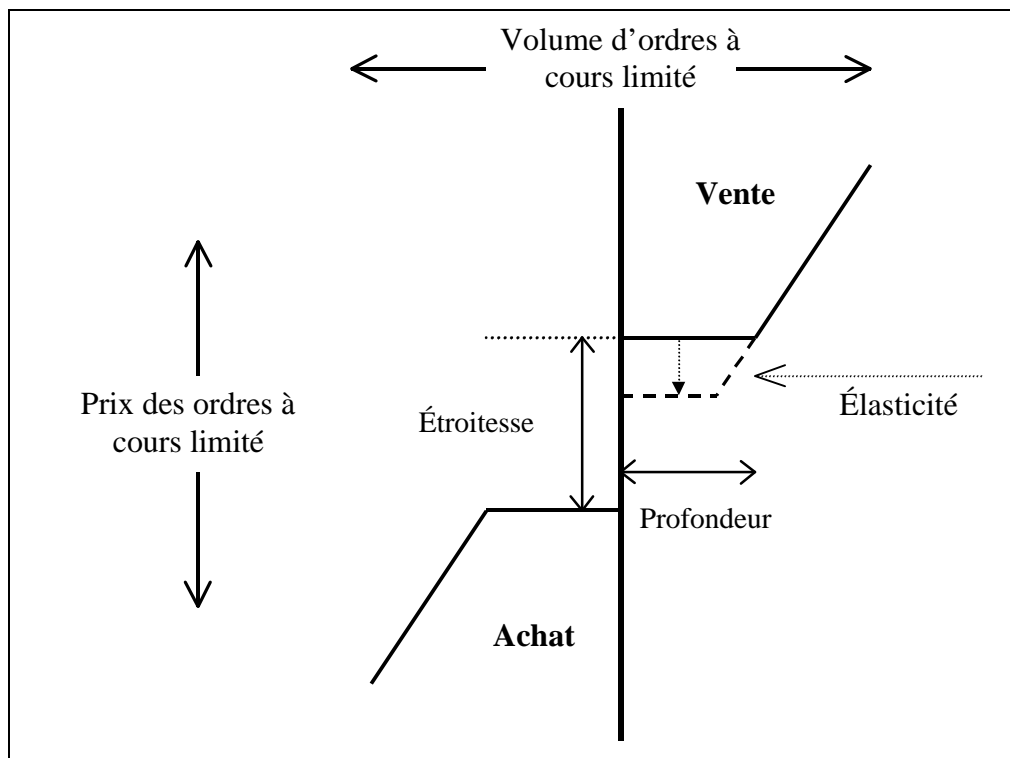
Graphiquement, comme le montre la figure 2.1, l'étroitesse peut être représentée par le degré de discontinuité existant entre la courbe d'offre et de demande au point d'origine. Ce dernier n'indique que le prix auquel peut s'attendre un acheteur ou un vendeur qui ne souhaite échanger qu'une petite quantité d'actions.

Une autre dimension de la liquidité est constituée par la profondeur qui est définie comme le volume pouvant être échangé à un niveau de prix défini. Les lignes horizontales partant

de l'origine correspondent au nombre maximum d'actions qui peut être échangé sans affecter le prix affiché. Selon Glosten et Harris (1988), la fourchette affichée sur un marché gouverné par les prix est bien souvent plus large que la fourchette cotée, par conséquent, il est probable selon eux que la profondeur réelle soit plus importante que celle affichée. Il n'en demeure pas moins que le principe reste le même en particulier sur les marchés gouvernés par les ordres où la fourchette cotée correspond à la fourchette affichée ; en revanche, il est vrai que la profondeur réelle est nettement plus importante en raison des ordres cachés.

FIGURE 2.1

Trois concepts de la liquidité



Source : Engle et Lange (1997)

Enfin, le dernier concept correspond, tel que le décrit la figure 2.1, à l'élasticité prix-volume : plus ce dernier est élevé, plus le prix obtenu va dévier de sa meilleure limite affichée. L'écart de prix dépend de la quantité des ordres à cours limite en attente. Une courbe à forte pente est le signe d'une faible liquidité. La figure 2.1, certes simplificatrice, possède le mérite de clarifier ces trois concepts de liquidité et de bien mettre en relief l'importance du temps : plus l'investisseur est patient, meilleur est le prix qu'il obtient.

L'approche naturelle pour estimer l'inclinaison de la pente des courbes d'offre et de demande est de mesurer le volume net et la variation de prix sur un intervalle de temps fixé. La variation de prix par action de la demande en excès est alors une estimation de l'élasticité. Il existe plusieurs raisons pour ne pas suivre ce raisonnement selon les auteurs : L'excès de demande peut être positif ou négatif, et selon la base de données possédée, il peut être particulièrement difficile d'affecter un signe aux volumes échangés sur l'initiative d'un acheteur ou d'un vendeur. La possibilité de diviser par un nombre proche de zéro est élevée ; par conséquent, il est prévisible de devoir faire face à des discontinuités. Si l'intervalle de temps retenu est court, les problèmes sont donc amplifiés. En outre, la discrétisation des prix signifie qu'il y a très peu de valeurs possibles au numérateur à pouvoir être anticipées. Le choix d'un intervalle relativement long n'est pas sans défaut car il deviendrait possible de déterminer une dynamique d'offre et de demande.

Engle et Lange (1997) proposent donc une alternative pour résoudre ces problèmes. Tout d'abord, l'intervalle de temps retenu est une « durée-prix » qui correspond au temps nécessaire pour que le prix coté varie d'un montant préalablement fixé. Le nombre de « durées-prix » est lié à la fois aux événements informationnels qui arrivent sur le marché ainsi qu'à l'écart de cours retenu. La détermination de cet écart va permettre d'ajuster le nombre « d'événements-prix ». La longueur du temps d'une durée prix va mesurer la vitesse à laquelle l'excès de demande peut être absorbé par le marché, or cette valeur attendue de la « durée-prix » est estimée comme étant inversement proportionnelle à la volatilité attendue par Engle et Russell (1998).

La VNET à travers la « durée-prix » s'apparente à une nouvelle mesure de la profondeur. Tant que la « durée-prix » correspond à une même variation de prix, la discrétisation du prix ne pose aucun problème statistiquement. En choisissant la profondeur comme étant le critère à mesurer, la VNET par prix devient une variable dépendante, et les propriétés statistiques sont alors suffisantes.

I.2.2 La formalisation de la liquidité à l'aide de l'indicateur VNET

Engle et Lange (1997) proposent donc d'analyser la profondeur du marché à travers une statistique nommée VNET. Cette statistique correspond précisément à l'excès de demande²² qui peut être échangé sans provoquer de variation de prix jugée significative. Cette statistique est formalisée de la manière suivante :

$$VNET = Ln \left| \sum_i (d_i * vol_i) \right| \quad (2.13)$$

Dans l'équation ci-dessus, d_i correspond à la variable indicatrice représentant la direction de la transaction.

Si l'échange est effectué sur l'initiative d'un acheteur, alors $d_i = +1$

Dans le cas où la transaction aurait été initiée par un vendeur, alors $d_i = -1$

Vol_i est le volume échangé lors de la transaction i pendant une « durée-prix » donnée.

Une telle mesure répond apparemment à la définition de la liquidité puisqu'elle intègre ses trois principales facettes : l'aspect volume, l'aspect temps et l'aspect prix. Nous pouvons dès à présent suggérer une amélioration ; il conviendrait de considérer différents états selon que la variation de prix est positive ou négative.

Une fois définie la statistique VNET, Engle et Lange (1997) cherchent à expliquer l'évolution durant la séance de cotation de cette statistique par des variables de microstructure. C'est pourquoi ils mettent en œuvre une régression classique. Engle et Lange (1997) proposent ainsi d'intégrer la série de variables suivantes²³ dans leur modèle explicatif :

◆ **La taille de la fourchette lors de l'événement prix précédent**, notée SP : la présence de cette variable (déjà présente à travers l'estimation de l'espérance de la « durée prix ») se

²² Au cours de la présentation, nous utilisons régulièrement le terme « excès de demande ». La demande correspond ici au désir d'échanger. Il peut aussi bien s'agir d'une offre (vente) que d'une demande (achat) d'actions.

²³ Les auteurs considèrent les transformées logarithmiques de ces variables.

justifie par le fait qu'elle est censée caractériser l'asymétrie d'information. Ainsi, une large fourchette suppose une plus grande variabilité du cours de l'action. Par conséquent, la profondeur estimée par la VNET, devrait être d'autant plus importante que la fourchette est étroite.

◆ **Le nombre de transactions pendant la « durée-prix » précédente**, noté *NB* : cette variable caractérise l'intensité des échanges. Un nombre élevé de transactions laisse supposer la présence d'investisseurs informés sur le marché. La profondeur devrait donc être d'autant plus faible que les initiés sont actifs sur le marché. D'une part, leur présence dégarrit le carnet et, d'autre part, les investisseurs informés limitent l'apparition des offreurs de liquidité.

◆ **Le volume de transactions pendant la durée prix précédente**, noté *VOL* : Le volume peut lui aussi être associé à l'intensité des échanges. La probabilité d'une corrélation avec le nombre de transactions est plus qu'élévée et risque de nuire au modèle. Ceci étant, le volume peut refléter aussi la qualité des informations transmises par l'intermédiaire de la fréquence d'appariements, et constituer ainsi un apport à la modélisation de la VNET.

◆ **Le nombre de transactions pendant la « durée-prix » précédente lorsque la fourchette est jugée large**, noté *NBLSP* : Engle et Russell (1994) indiquent que l'estimation du modèle ACD est meilleure lorsque l'on effectue une distinction selon la largeur du *Bid/Ask spread*. L'intensité des transactions est plus significative lorsque la fourchette est considérée comme importante. Ceci justifie certainement la présence de cette variable.

◆ **Le volume de transactions pendant la « durée-prix » précédente lorsque la fourchette est jugée large**, noté *VOLLSP* : pour les mêmes motifs que la variable précédente, Engle et Lange (1997) effectuent une distinction selon le niveau de la fourchette.

◆ **L'écart de prix justifiant le nouvel événement de prix**, noté *ECP* : cette variable se justifie par le fait que la VNET doit être d'autant plus importante que la variation est jugée significativement élevée.

◆ **L'espérance conditionnelle de la durée**, notée *ESPT* : l'introduction de cette variable est logique dans la mesure où la statistique VNET correspond au volume en excès durant

cette période estimée. La question qui se pose est de savoir si une « durée-prix » longue (le marché est peu volatil mais aussi certainement moins nourri en termes d'activité) signifie que c'est une période très liquide qui autorise des déséquilibres en termes de volumes signés. Les auteurs suggèrent d'estimer l'espérance conditionnelle de la « durée-prix » pour chaque titre à l'aide d'un modèle WACD(1,1) auquel on rajoute le niveau de la fourchette décalé d'un retard, comme formulé ci-dessous.

$$y_i = w + a * \tilde{X}_{i-1} + b * y_{i-1} + dSP_{i-1}$$

Avec \tilde{X}_i , la durée désaisonnalisée qui correspond au temps séparant le $i^{\text{ème}}$ et le $(i-1)^{\text{ème}}$ événement-prix.

où y_i n'est autre que l'espérance conditionnelle de la durée sachant son passé.

et SP_i correspond à la fourchette lors de l' $i^{\text{ème}}$ événement prix.

Ce modèle de prévision apparaît, selon Engle et Lange (1997), comme étant le meilleur estimateur sur la période d'étude considérée.

♦ **Le rapport entre la « durée-prix » observée et son espérance (l'erreur de prévision)**, notée *ESPTR* : cette variable exprime l'erreur de prévision ; la « durée prix » est plus ou moins longue que prévue, ce qui signifie aussi que le marché est plus ou moins volatil qu'estimé. Une erreur d'estimation peut ainsi correspondre à une augmentation de la fréquence des échanges, c'est-à-dire à l'arrivée d'une information sur le marché.

I.2.3 Une première estimation de VNET

Engle et Lange (1997) effectuent tout d'abord plusieurs simulations à partir de titres individuels pour aboutir à une estimation groupée²⁴ qui leur permet d'obtenir des coefficients significatifs, ce qui n'était pas le cas lors de l'analyse de la régression action par action.

Le résultat global est le suivant :

$$VNET = -0,33 \cdot SP_{t-1} + 0,15 \cdot VOL_{t-1} - 0,15 \cdot NB_{t-1} + 0,39 \cdot ESPT_t + 0,38 \cdot REPST_t$$

²⁴ Une estimation effectuée à partir de données empilées.

Le premier commentaire concerne l'absence de certaines variables. On peut noter que la variables d'échange (i.e. le nombre de transactions et le volume), et que la variable de saut de prix n'ont pas été retenues. Ceci peut être expliqué par les motifs suivants : tout d'abord, en ce qui concerne les variables d'activité, les auteurs n'ont pas effectué de distinction réelle entre les deux régimes de fourchette ; il aurait peut-être été, préférable d'introduire tout simplement le cas où la fourchette était étroite. En ce qui concerne le saut de prix, son absence peut s'expliquer par le fait que cette variable correspond très certainement au seuil d'événement-prix dans une très large majorité des cas. Par conséquent, nous avons trop peu « d'évènements-prix » différents pour révéler un coefficient significatif. Si les données avaient été empilées en considérant différents seuils d'événement, on peut penser que la variable aurait été largement significative.

Le second commentaire porte sur les signes des relations entre la profondeur, mesurée par la statistique VNET, et les variables explicatives.

On obtient une relation négative avec la fourchette de la « durée prix » précédente. La fourchette possède aussi un impact sur la VNET indirectement à travers l'espérance de la « durée-prix » calculée par le modèle de durée conditionnelle. Cette relation négative n'est pas surprenante. En effet, la fourchette et la profondeur correspondent à deux facettes de la liquidité d'où la corrélation partielle. Une large fourchette est associée à une faible profondeur. De plus, une large fourchette laisse présager des variations de prix plus importantes donc non seulement d'une « durée-prix » plus faible mais aussi une demande de volume en excès plus faible.

La relation avec l'intensité des échanges mesurée par le nombre de transactions par « durée prix » corrobore les prédictions ayant pour fondement les différentes théories relatives à l'asymétrie d'information. Les modèles d'asymétrie d'information prévoient un coefficient négatif pour le nombre de transactions si la concentration des échanges résulte d'un flux d'informations.

La relation positive avec les volumes de transactions est quelque peu gênante dans la mesure où le volume reflète aussi l'activité dont on vient de constater le lien négatif avec la VNET. Une première explication peut trouver son origine dans le fait que Jones, Kaul et Lipson (1993) estiment que le volume est endogène au nombre de transactions et que par conséquent le signe de la relation est dû à un problème statistique. Une autre explication

fournie par Engle et Lange (1997) tient au fait que la VNET correspond à un montant de volume net. Il est alors probable d'obtenir une VNET élevée si le volume global échangé durant la « durée-prix » est lui aussi très important. Un volume élevé implique en effet un faible pourcentage de déséquilibre dans les ordres. Le coefficient obtenu est significativement inférieur à l'unité. Ceci indique que l'accroissement du volume prévoit une augmentation proportionnellement moindre de la profondeur, car il accroît la probabilité que les échanges additionnels soient à l'initiative des informés.

Il existe une corrélation positive entre la VNET et l'espérance de la durée prix. Comme nous l'avons exprimé lors de la présentation des variables, l'espérance conditionnelle de la durée doit être considérée comme une fonction réciproque de la volatilité. Or nous savons que la volatilité est d'autant plus faible que le marché est profond : ceci justifie pleinement le signe positif de la relation mise en relief d'autant plus que la volatilité est associée à l'arrivée d'informations sur le marché et ainsi à la présence potentielle d'investisseurs informés.

La dernière variable très significative est l'erreur de prévision de la « durée-prix » ; On obtient une relation positive là encore. Une erreur de prix positive correspond à une durée observée supérieure à la durée calculée par le modèle ACD. Par conséquent, le marché s'est montré moins volatil que prévu. Ainsi, pour ces motifs, la profondeur devrait être plus élevée. Dans le cas inverse, une durée plus courte que prévue signifie que la volatilité est plus forte en raison de l'arrivée d'une information sur le marché d'où la modification plus rapide des cours. Ainsi, Engle et Lange (1997) considèrent qu'à l'ouverture du marché, le teneur de marché (i.e. le *specialist*) s'attend à ce que le flux d'informations arrivant sur le marché lui dicte les mouvements de prix d'où un accroissement de sa fourchette. Finalement, si la révélation des informations est moindre que prévue, le teneur de marché est peu motivé pour réviser ses prix d'où une incitation à offrir de la liquidité et un délai plus élevé entre les événements-prix. Le marché affiche alors une meilleure profondeur.

Selon les mêmes auteurs, un intervenant financier peut influencer cette erreur de prévision en effectuant des transactions rapides. Ainsi, l'impatience de l'investisseur est prise en compte par le marché comme une nouvelle. En conséquence, les coefficients²⁵ uniformément positifs sur ce choc temporel reflètent les coûts associés avec des échanges

²⁵ Le coefficient est le seul à être significatif sur l'ensemble des titres individuels.

instantanés. La profondeur du marché est réduite, et le volume qui peut être négocié avec un impact sur un prix déterminé s'en trouve réduit. Les résultats de la régression indiquent que le coefficient prend une valeur peu différente de 0,4 ce qui signifie que l'investisseur qui pourrait se montrer deux fois plus patient ferait face à un marché plus profond de 40%.

Du point de vue du gestionnaire, les auteurs proposent une simulation, en supposant que l'espérance de la « durée-prix » est sous le contrôle de l'investisseur. Celui-ci veut savoir combien d'actions il peut échanger à une date avec un prix connu à l'avance. La réponse à cette ambitieuse question est la prévision de la VNET conditionnellement à l'espérance conditionnelle de la « durée-prix ».

I.2.4 Analyse de la stabilité du modèle VNET

Lange (1998) examine la robustesse de la modélisation VNET, et étudie son utilité pour prévoir et mesurer l'évolution en séance de la profondeur du marché. Pour cela, il effectue de nouvelles estimations des coefficients sur le second semestre 1997, ce qui lui permet d'ailleurs d'observer le comportement de la statistique durant la séance du mini-krach qui s'est déroulé le 27 octobre 1997.

L'analyse est rendue difficile dans la mesure où la première période d'étude s'étale sur la fin de l'année 1990 et le début de l'année 1991. En effet, on considère l'estimation de la VNET à chaque événement-prix ; ceux-ci sont déterminés à partir d'une variation fixée au préalable. Or ces écarts de cours nominaux ne représentent plus la même variation relative. Ainsi, alors que ces seuils permettaient de constater entre dix et quinze « durées-prix » par séance sur la première période d'étude, ce nombre franchit allègrement le cap des trente durées pour certaines actions de l'échantillon. En outre, les règles de cotation sur le NYSE, lieu de l'étude, ont été modifiées dans la mesure où la variation de prix minimale, i.e. le *tick*, est désormais de 1/16 *dollar* contre 1/8 *dollar* auparavant. Il est donc nécessaire à l'auteur d'effectuer quelques ajustements. De plus, le contexte général du marché est bien différent puisque la rentabilité mensuelle du NYSE est de 0,3% sur la présente étude contre 4,1% lors de la première étude. Ce contexte différent peut paraître gênant pour vérifier la stabilité de la modélisation, mais il doit être considéré ici comme une période adéquate pour prouver la robustesse de la mesure statistique VNET.

En fait, du point de vue de Lange (1998), le principal problème est d'ordre économétrique puisque le calcul de l'espérance conditionnelle de la durée entre les évènements de prix à l'aide de la même modélisation [WACD(1,1) avec la fourchette comme variable additionnelle] ne semble pas être fiable pour l'ensemble des actions de l'échantillon. En effet, l'autocorrélation des « durées-prix » n'est pas capturée totalement par le modèle. C'est pourquoi Lange (1998) est amené à procéder là encore à quelques restrictions sur l'estimation de la durée afin de ne pas donner un poids trop important à l'erreur de prévision de la « durée prix » dans l'équation de la VNET au détriment de l'estimation conditionnelle de cette même « durée-prix ».

Le constat d'ordre général, concerne les évolutions moyennes respectives de la VNET, des volumes de transaction et de la volatilité entre les deux périodes d'étude. Lange (1998) constate que la profondeur de marché mesurée par VNET s'est améliorée pour une large majorité des actions de son échantillon. VNET mesure la profondeur en unité de volumes par variation de prix. Par conséquent, on doit s'attendre à ce que la VNET moyenne augmente conjointement avec les volumes de transaction tandis que le niveau de volatilité décroît, d'après le sens des relations établies lors de la première étude. Or ce résultat est vérifié puisque Lange (1998) remarque que les actions dont la VNET a baissé, présentent une volatilité accrue et une faible augmentation des volumes échangés. On obtient la réciproque si on considère les actions dont la VNET s'est élevée.

En ce qui concerne l'examen approfondi titre par titre, Lange (1998) démontre tout d'abord que les relations entre la statistique VNET et ses variables explicatives sont préservées. En revanche, pour ce qui est de la stabilité des coefficients, les résultats sont plus mitigés, puisqu'une fois corrigés des problèmes d'espérance conditionnelle de « durée-prix » ; un tiers des actions de l'échantillon ne se conforme pas à l'hypothèse de stabilité des coefficients.

Lange (1998) considère, que de tels résultats²⁶ sont encourageants dans la mesure où l'on compare un échantillon relativement restreint (16 actions) sur deux périodes d'étude très éloignées. En outre, il se pose le problème de la spécification de l'espérance de la « durée-prix ». Il faut aussi remarquer qu'un coefficient qui diffère sur une période de sept ans n'a

²⁶ Il convient de se souvenir que la modélisation initiale titre par titre ne donnait pas entière satisfaction.

rien d'anormal, il suffit de considérer l'exemple du coefficient bêta des actions qui est loin d'être stable dans le temps.

I.2.5 Influence de l'événement prix sur la VNET

L'analyse de la robustesse de la statistique VNET a permis de mettre en avant le problème que posait la détermination de la largeur de la fenêtre d'agrégation, i.e. la taille du seuil de déclenchement au-delà duquel nous estimons qu'il y a réalisation d'un « événement-prix ». Engle et Russell (1997) considèrent qu'il faut retenir la médiane de la taille de la fourchette tandis qu'Engle et Lange (1997) semblent retenir un seuil permettant d'obtenir un certain nombre de « durée-prix » par séance. Considérons la méthodologie d'Engle et Russell (1997), si la médiane de la taille de la fourchette équivaut à X échelons de cotation et si cet écart de cours correspond à une variation de cours de 0,20%, peut-on considérer un tel écart comme un « événement-prix » significatif ? Du point de vue du titre financier en lui-même, il s'agit effectivement d'un « événement-prix » mais pour l'investisseur, soumis à des coûts de transaction, qui souhaite utiliser la « durée-prix » comme un instrument de gestion, la réponse est plus mitigée. En revanche, 0,20% pour un gérant de fonds, c'est déjà un gain très appréciable au vu des fonds qu'il doit gérer et ce d'autant plus qu'il doit faire face à une forte concurrence.

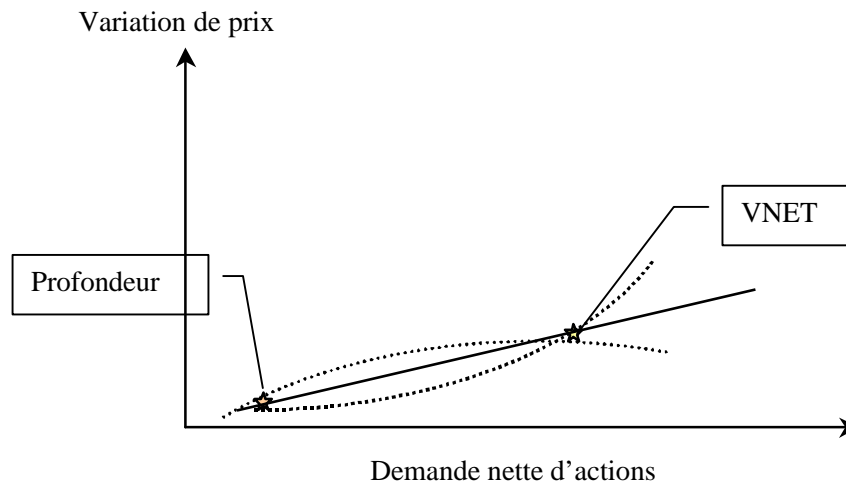
Lange (1998) analyse l'incidence du choix de cet écart de cours. Comme nous l'avons signalé précédemment, la statistique VNET illustre la profondeur dynamique du marché mais peut aussi être interprétée comme la pente d'une courbe de réaction du marché. Avec simplement une mesure de la VNET, il est possible de déduire la pente moyenne de cette courbe de réaction mais non son profil. La figure 2.2 illustre notre propos ; une fonction linéaire, concave ou encore convexe peut être estimée à partir de deux points (les volumes associés à la meilleure limite et la VNET).

Si on détermine plusieurs seuils de déclenchement à partir duquel on définit un « événement-prix », on est susceptible de pouvoir estimer d'autres points pour construire de manière plus précise cette courbe de réaction du marché. Lange (1998) simule ainsi pour chaque événement prix, la VNET correspondante. Ainsi plus la taille du seuil de déclenchement est grande, plus les durées prix, moins nombreuses, sont longues. Comme l'indiquent les figures 2.3, la statistique VNET est une fonction croissante de la taille de la

variation de prix autorisée. La courbe qui présente une faible (forte) pente représente une action liquide (illiquide).

FIGURE 2.2

Courbes implicites de réaction de marché

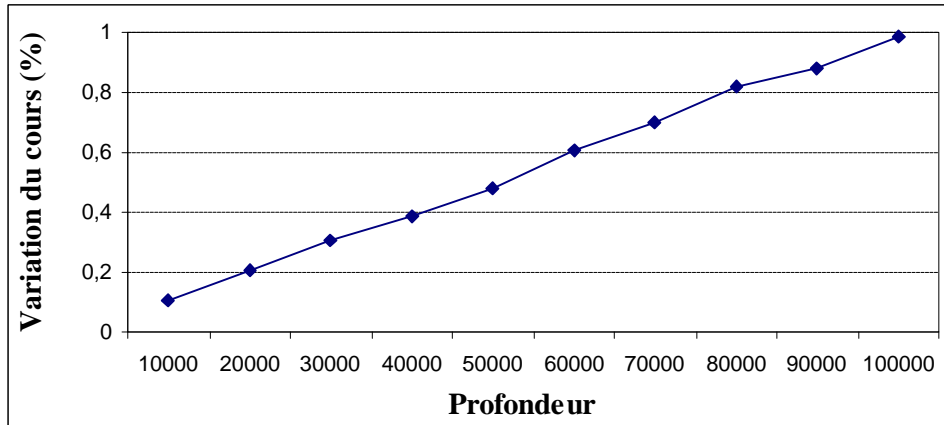


Source : Lange (1998)

En outre, Lange (1998) s'aperçoit que cette courbe de réaction dynamique du marché tend à être linéaire pour une très large majorité des actions de son échantillon [figures 2.3]. De plus, on constate que le premier point de la fonction linéaire correspond à la profondeur de la meilleure limite du carnet d'ordres. On considère que ce volume, par hypothèse, est associé à la variation minimale de cours, i.e. le *tick*. Le fait que ce point se situe sur cette droite tracée à partir des différentes observations de la statistique VNET, tend à soutenir l'idée selon laquelle la statistique est bien une mesure appropriée de liquidité.

FIGURE 2.3A

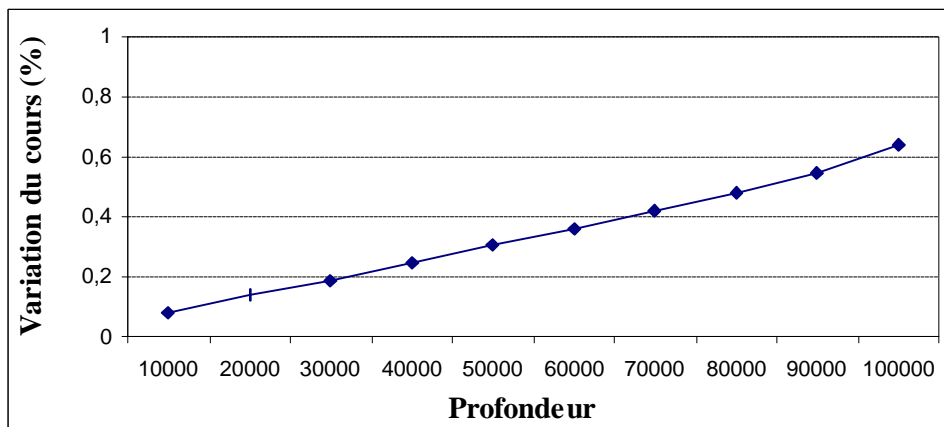
Courbe moyenne de réaction de marché (simulation)
Action peu liquide



Source : Lange (1998)

FIGURE 2.3B

Courbe moyenne de réaction de marché (simulation)
Action très liquide



Source : Lange (1998)

Il paraît techniquement possible de modéliser et donc de prévoir la courbe de réaction du marché durant la séance. Lange (1998) propose alors de modéliser à partir de la régression initiale chaque VNET selon son seuil de déclenchement. Ainsi, les simulations ci-dessus (figures 2.3A et 2.3B) représentent la construction de deux courbes à partir de l'estimation de la VNET pour neuf seuils de déclenchement. Les coefficients diffèrent quelque peu (en particulier la constante) selon le niveau d'écart de cours autorisé, mais demeurent

relativement stables statistiquement lors d'un changement de seuils. Toutefois, Lange (1998) démontre, à l'aide d'un test joint, que les modélisations sont significativement différentes. L'inconvénient majeur de cette série de modélisations concerne la significativité de la fourchette²⁷ dans le modèle de régression. Cependant Lange (1998) montre, à partir de nouvelles simulations, que cela n'a pas de conséquences importantes sur la prévision de la profondeur dynamique ; de plus, il convient de ne pas oublier que la fourchette est déjà présente dans la régression à travers l'espérance conditionnelle de la « durée-prix ». En ce qui concerne les autres variables, les relations établies précédemment sont confirmées à savoir que quel que soit l'événement prix retenu, les corrélations demeurent négatives avec le nombre de transactions et la volatilité du marché, et positive avec le volume d'actions échangés.

I.2.6 Analyse de la performance de la VNET à mesurer la liquidité en séance

Nous avons présenté une nouvelle mesure de liquidité - VNET - introduite par Engle et Lange (1997). Cette mesure est directement liée à la volatilité du marché grâce à un changement de temps déterminé par la « durée-prix ». De fait, si deux événements rapides surviennent à quelques instants d'intervalle - un premier événement provoque une augmentation du prix de trois échelons et un second ramène le cours à son niveau initial -, une approche calendaire enregistre une variation de cours nulle durant la période de temps analysée. En revanche, la méthode de la « durée-prix » va générer deux observations de la statistique VNET - une après chaque événement prix. Ainsi, cette nouvelle mesure permet de délivrer une image plus juste de l'évolution de la liquidité durant cette période. Le principal avantage de cette mesure est qu'il semble possible de la modéliser et donc de la prévoir.

Lange (1998) teste son modèle sur les séances des 27 et 28 octobre 1997, séances qui se caractérisent par respectivement une forte chute du cours des actions et une remontée tout aussi importante. Ces séances de tempête boursière sont vraiment particulières en matière d'évaluation de la liquidité car elles se caractérisent par un fort volume d'activité, impliquant une hausse de la liquidité, et une forte volatilité, générant une baisse de la

²⁷ La fourchette n'est pas significative pour certains seuils ; de plus, le signe de son coefficient n'est pas

liquidité. Par conséquent comment juger de l'état du marché : de nombreuses mesures indiquent une baisse de la liquidité (fourchette, lambda, volatilité...) tandis que d'autres indiquent une hausse (volume d'activité). Une telle période d'étude paraît ici idéale tant les évolutions sont contrastées et tant les mesures de liquidité sont mises à l'épreuve.

Les résultats de Lange (1998) sont les suivants : tout d'abord l'auteur dresse le bilan de ces deux journées de bourse. Il constate, que les durées entre transactions sont plus rapides que de coutume, les fourchettes sont plus larges, le niveau des volumes et la taille moyenne des ordres sont plus élevés ; il en est de même bien évidemment pour la volatilité du marché, i.e. les « durées-prix » sont plus nombreuses et plus courtes. La statistique comprenant ces différents paramètres dans sa modélisation, une incertitude se fait jour. Lange (1998) compare donc la VNET réalisée à la VNET prévue par le modèle. Il s'aperçoit que si nous devons comparer ces deux séances à une période de « stabilité », la profondeur réalisée aussi bien que celle prévue, ne dévient que très faiblement du niveau qu'elles ont habituellement, contrairement à l'activité boursière et à la volatilité.

En fait, la relative stabilité du niveau de liquidité mesurée par la VNET, un indicateur synthétique, est la conséquence d'un équilibre entre les facteurs influençant positivement et négativement le critère de liquidité. Ainsi, bien que très peu de transactions soient absorbées dans une « durée-prix », le large volume de chaque transaction fait rester la mesure VNET à un niveau élevé. Lange (1998) détecte cependant une certaine asymétrie entre les deux séances. Les prix ont apparemment baissé plus rapidement lors du krach qu'ils n'ont remonté le lendemain dans une séance beaucoup plus active. Lors de la séance de baisse, la liquidité réalisée a eu tendance à sous-performer sa prévision (11 actions sur 16). Le lendemain, alors que le marché attendait un rebond, la profondeur réalisée s'est révélée être supérieure à celle anticipée (14/16). Selon Lange (1998) et les prévisions du modèle, le marché a montré un bon niveau de liquidité lors du Krach. Le lendemain, quand le marché s'est repris, la profondeur s'est révélée élevée de manière non prévue, ce qui a engendré un nombre d'achats plus important que d'habitude d'où le soutien des cours.

Finalement, Lange (1998) teste la stabilité des coefficients sur ces deux séances particulières ; il trouve que la valeur de ceux-ci n'a pas été significativement modifiée. Ainsi, Lange conclut à la robustesse et à la significativité du modèle de régression VNET.

constant.

Ceci signifie que la modélisation VNET n'est pas altérée à travers le temps, quel que soit le seuil de déclenchement des événements prix retenu et enfin quelles que soient les conditions de marché.

La mesure synthétique de liquidité proposée par Engle et Lange (1997) apparaît difficile à mettre en œuvre rapidement néanmoins un tel indicateur présente un réel intérêt car il offre une nouvelle vision de la liquidité sous un aspect dynamique jusque là ignoré. Un tel estimateur paraît donc être en mesure d'aider les investisseurs dans leur décision d'effectuer des transactions dans les meilleures conditions.

SECTION II - ANALYSE EMPIRIQUE

L'objectif de cette section²⁸ est de mettre en œuvre une statistique VNET adaptée au marché français. Pour cela nous allons procéder en trois étapes. Tout d'abord, nous allons déterminer le modèle de durée qui permet de « capturer » l'autocorrélation des durées qui séparent les événements prix. On va tenter par la suite d'améliorer ce modèle en intégrant des variables additionnelles. Enfin, à partir du meilleur modèle de durée retenu, on va rechercher les variables explicatives qui permettent de modéliser la statistique VNET.

Ainsi cette section se décompose comme suit : un premier paragraphe introductif aux modélisations économétriques présente les données utilisées et quelques éléments méthodologiques ; le deuxième paragraphe permet de mettre en relief la présence d'une saisonnalité et d'une autocorrélation entre les durées des événements prix ; le troisième paragraphe est destiné à apprécier le choix du modèle de durée ; le suivant a pour objectif d'améliorer le modèle de durée initialement retenu ; le cinquième et dernier paragraphe conclut par la mise en œuvre d'un modèle de régression VNET.

II.1 ÉLÉMENTS MÉTHODOLOGIQUES

L'objectif est de présenter les différents éléments méthodologiques nécessaires à l'analyse économétrique et empirique. Dans un premier temps, nous exposons les données et la période d'étude retenues. Par la suite, nous présentons la définition de l'événement prix. Enfin, nous concluons par l'estimation de l'intensité des événements prix.

II.1.1 Présentation des données

Afin de mettre en œuvre notre analyse, nous avons utilisé les données à haute-fréquence issues de la base BDM de ParisBourse SBF. Il s'agit des bases de données des meilleures limites et des transactions observées sur les actions Pernod-Ricard et Saint-Gobain. Notre période d'étude s'étale du 25 janvier 1999 au 24 mai 1999, ce qui correspond à quatre mois boursiers. Il convient de signaler que nous avons ôté les séances de dénouement des

contrats à terme qui présentent un profil d'activité différent des séances traditionnelles [Le Saout (1999a)]. Ceci permet d'obtenir une estimation plus fiable des fonctions *splines*.

Dans la littérature relative à la modélisation des durées entre évènements de cotation, les différents auteurs [Bauwens et Giot (1998), Engle et Russell (1999)...] ont fondé leur analyse à l'aide du milieu de la fourchette. Pour notre part, nous avons opté pour une distinction entre les flux d'achat et de vente réalisés à partir du croisement des tables des meilleures limites et des transactions. L'origine de ce choix est liée à l'interprétation d'un mouvement de la fourchette. Nous pouvons observer des mouvements de prix significatifs en raison aussi bien d'une transaction, de l'annulation d'un ordre ou encore du placement d'un ordre à cours limité. De plus, le milieu de la fourchette ne prend pas en compte les volumes associés aux meilleures limites, ce qui signifie qu'un ordre de faible taille possède autant d'incidence qu'un ordre de taille importante²⁹.

II.1.2 Détermination de l'événement prix

La première étape de notre analyse de la durée entre différentes variations de cours consiste à déterminer le seuil au-delà duquel nous définissons l'existence d'un événement prix. Engle et Russell (1998) suggèrent de retenir la fourchette médiane. Nous avons donc décidé de retenir cette méthodologie.

Toutefois un problème s'est posé à nous. Les fourchettes médianes des titres Pernod-Ricard et Saint-Gobain diffèrent. Ainsi, elle s'élève à cinq *ticks* pour le premier titre et à trois *ticks* pour le second titre. Plus ennuyeux est le fait que la valeur de l'échelon de cotation n'est pas identique nominalement et relativement. En effet, l'échelon de cotation de l'action est de 0.05 euro pour Pernod-Ricard tandis qu'il est de 0.1 euro pour Saint-Gobain, soit respectivement 0.087% et 0.067% de leur cours moyen. Par conséquent, le seuil d'événement prix est de 0.43% pour Pernod-Ricard et seulement de 0.20% pour Saint-Gobain. Peut-on juger de telles variations comme des écarts de prix significatifs ?

²⁸ Nous tenons à remercier vivement Pierre Giot et Sophie Coutant. La réalisation des études empiriques sous GAUSS de cette section n'aurait pu voir le jour sans leur aide.

²⁹ A titre indicatif, nous avons comptabilisé les mouvements du barycentre de la fourchette qui lui tient compte des volumes. Nous avons constaté près de trois fois plus de mouvements significatifs de prix qu'en retenant le prix d'une transaction effectivement réalisée.

Face à ce dilemme, et afin de pouvoir effectuer une tentative de comparaison entre les deux actions, nous avons décidé de définir deux montants de seuil de variation : un écart de trois *ticks* et de cinq *ticks* pour chacune des actions. Nous n'avons pas retenu de seuil de prix plus élevé. Ce choix s'explique par le fait que, d'une part, nous avons respecté la méthodologie initiale d'Engle et Russell (1994), et que d'autre part, si nous avions retenu un niveau de 1%, nous aurions fait face à un autre problème : lors de certaines séances peu volatiles, il y aurait eu trop peu d'évènements-prix pour pouvoir tenter de modéliser les durées conditionnelles dont le caractère autorégressif était quasiment nul³⁰.

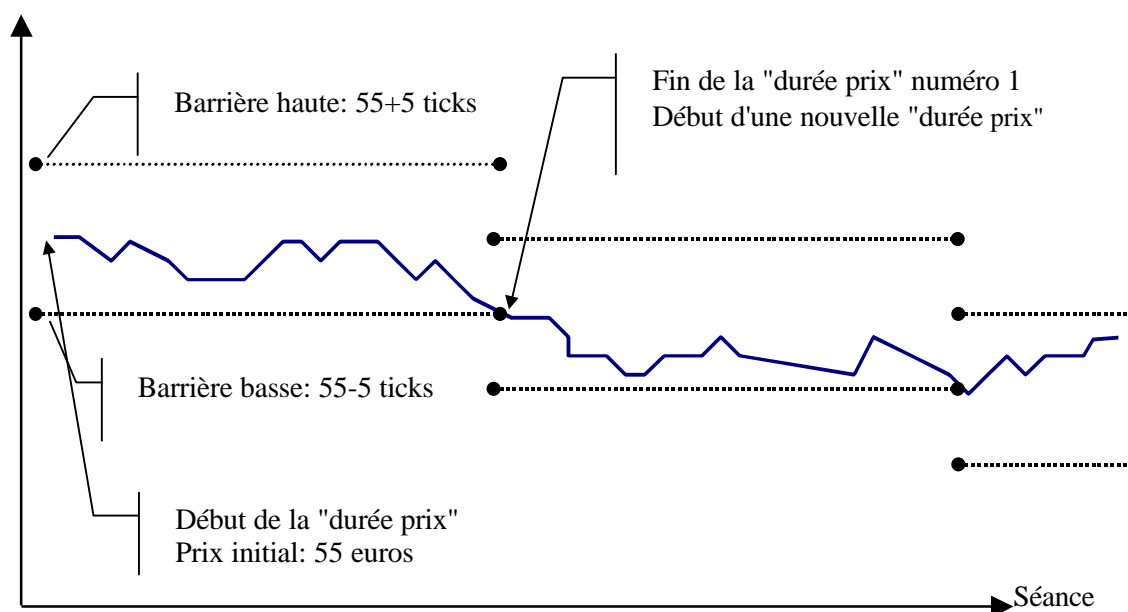
Si on considère l'exemple de l'action Pernod-Ricard dont on juge la variation de prix significative lorsque l'on dépasse le seuil de cinq échelons de cotation. On suppose un cours de 55 euros à l'instant t qui marque le début d'une durée prix. Cette durée prix prendra fin lorsque le cours touchera soit le cours plancher de 54,75 ($55-5*0,05$) euros soit le cours plafond de 55,25 ($55+5*0,05$) euros. La figure 2.4 reprend cet exemple: la fin de la durée prix initiale marque le début d'une nouvelle durée prix. De nouveaux seuils sont fixés, ici $54,75+0,25$ euros et $54,75-0,25$ euros.

Nous avons donc appliqué ce traitement à nos données. Le tableau 2.1 présente les caractéristiques des durées entre évènements-prix selon l'action, le sens de la transaction et le seuil de variation considérés. Nous constatons une fréquence d'apparition d'évènements-prix plus élevée pour l'action Saint-Gobain que pour l'action Pernod-Ricard. En outre, nous pouvons remarquer que le passage d'un seuil de trois *ticks* à cinq *ticks* permet de réduire de près d'un tiers le nombre d'évènements-prix pour l'action Pernod-Ricard contre un peu plus de quarante pour cent pour l'action Saint-Gobain.

³⁰ Une telle modélisation est possible mais elle exige une période d'étude nettement plus grande et un ordinateur dotée d'une très importante capacité mémoire.

FIGURE 2.4

Détermination d'un évènement prix


TABLEAU 2.1

Caractéristiques des durées entre les évènements prix

Pernod-Ricard				
Sens	Achat	Vente	Achat	Vente
Variation	3 ticks	3 ticks	5 ticks	5 ticks
Variation relative	0,26%	0,26%	0,43%	0,43%
Nombre d'événements	1579	1430	1099	989
Moyenne/séance	20,24	18,33	14,09	12,68
Durée Moyenne	1159	1251	1589	1737
Durée Médiane	542	602	773	812
Écart-type	1760	1836	2373	2599
Saint-Gobain				
Sens	Achat	Vente	Achat	Vente
Variation	3 ticks	3 ticks	5 ticks	5 ticks
Variation relative	0,20%	0,20%	0,34%	0,34%
Nombre d'événements	4089	4336	2376	2453
Moyenne/séance	52,42	55,59	30,46	31,45
Durée Moyenne	467	442	795	775
Durée Médiane	267	239	420	397
Écart-type	638	604	1091	1082

II.1.3 Analyse de l'intensité des évènements prix

Afin d'illustrer la fréquence des évènements prix, nous reprenons la méthodologie développée par Gouriéroux, Jasiak et Le Fol (1999) qu'ils appliquent à la fréquence des appariements. Le principe de ces travaux consiste à analyser le temps écoulé entre deux évènements de cotation consécutifs. Ceci peut être ainsi effectué à l'aide du modèle périodique suivant :

$$\text{Pr } ob_k(\Delta t) = \begin{cases} 0(\Delta t) & \text{si } k > 1 \\ \mathbf{I}(t)(\Delta t) + 0(\Delta t) & \text{si } k = 1 \\ 1 - \mathbf{I}(t)(\Delta t) + 0(\Delta t) & \text{si } k = 0 \end{cases}$$

où k indique le nombre d'évènements prix.

et où \mathbf{I} est égal au nombre moyen d'évènements-prix par unité de temps (\mathbf{I} est appelé aussi densité du processus poissonnien).

Nous supposons qu'il ne peut y avoir de durée nulle entre deux évènements prix. Cette hypothèse n'est absolument pas contraignante dans la mesure où une durée nulle correspond à un ordre qui touche plusieurs limites, or nous n'avons retenu que le cours résultant de la dernière séquence de transactions.

Le taux d'intensité des échanges peut être estimé à partir de méthodes non paramétriques. Dans le cas présent, il s'agit d'estimateurs à noyau ; ceux-ci présentent l'avantage d'éliminer les discontinuités.

Ainsi l'estimateur consistant du taux d'intensité est le suivant :

$$\hat{\mathbf{I}}(t) = \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^{N_m} \frac{1}{h} K \left[\frac{d_{n-1}(m) - t}{h} \right]$$

où h indique la longueur de la fenêtre qui tend vers zéro lorsque M tend vers l'infini.

M représente le nombre de séances de notre échantillon.

N_m dénomme le nombre d'évènements prix durant la séance m .

K est une fonction positive bornée appelé aussi noyau. Dans le cas de notre étude empirique, nous allons retenir le cas classique du noyau gaussien qui se définit de la manière suivante :

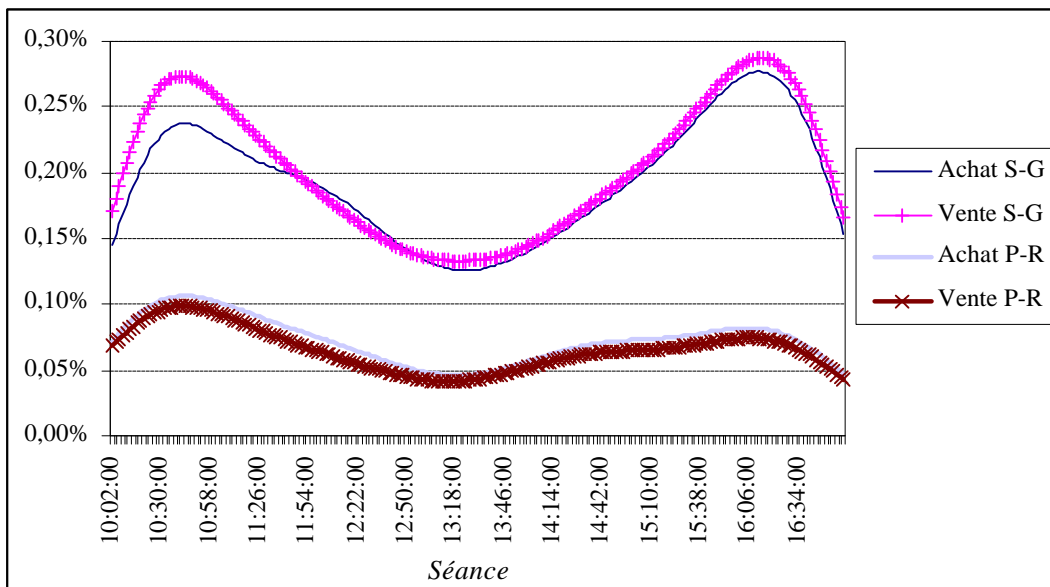
$$K(u) = \frac{1}{\sqrt{2\Pi}} e^{\left(-\frac{1}{2}u^2\right)}$$

Ainsi les graphiques 2.1 et 2.2 indiquent ainsi les taux d'événement prix, selon le seuil retenu, pour les actions Pernod-Ricard et Saint-Gobain.

Nous nous apercevons que les différentes courbes décrivent une forme de M à l'instar de la courbe de probabilité d'arrivée d'une transaction déterminée par Gouriéroux, Jasiak et Le Fol (1999). En outre, nous pouvons remarquer que les courbes ont tendance à se confondre. D'ailleurs la probabilité de non-événement pour chaque flux est quasiment semblable, même si cela est un peu moins évident pour l'action Pernod-Ricard. Ceci signifie que la fourchette est relativement stable dans le temps, et que dans le cas de Pernod-Ricard, c'est le flux d'achat qui semble diriger les prix dans la mesure où les courbes d'achat paraissent devancer les courbes de vente. Enfin, la volatilité de Saint-Gobain en cours de séance apparaît comme plus élevée que celle de Pernod-Ricard.

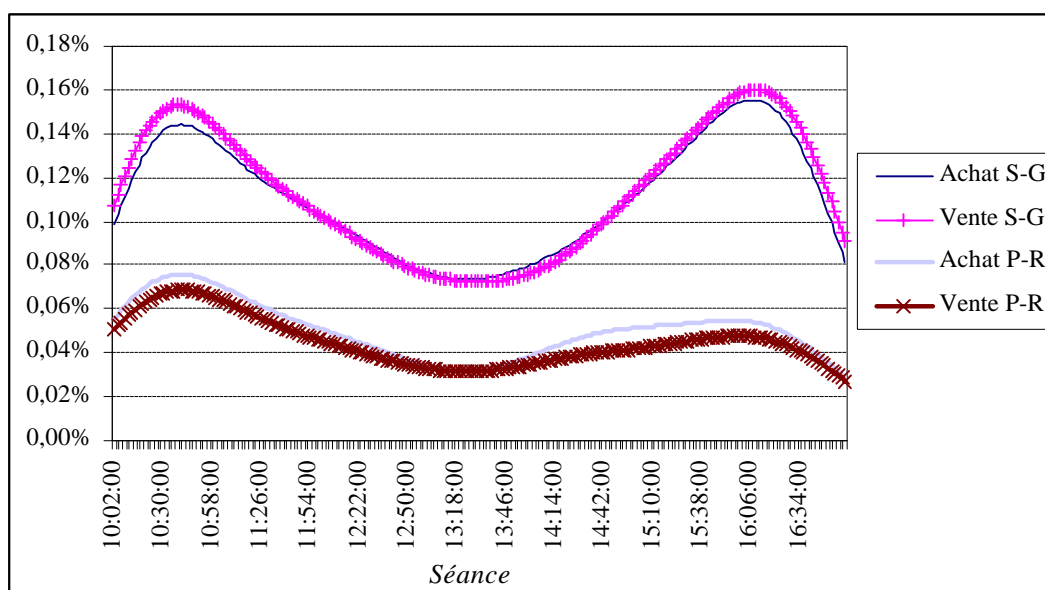
GRAPHIQUE 2.1

Probabilité d'arrivée d'un événement prix (Seuil: 3 ticks)



GRAPHIQUE 2.2

Probabilité d'arrivée d'un événement prix (Seuil: 5 ticks)



II.2 SAISONNALITÉS ET AUTOCORRÉLATION DES DURÉES ENTRE LES ÉVÈNEMENTS PRIX.

Nous allons dans un premier temps mettre en relief l'existence d'une saisonnalité intrajournalière et surtout présenter une méthode de désaisonnalisation. Dans un second temps, nous allons nous intéresser à la présence persistante d'autocorrélation entre les durées en dépit d'une désaisonnalisation. L'existence de cette autocorrélation est à l'origine de la recherche d'une modélisation adéquate qui fait l'objet des développements qui suivent.

II.2.1 Analyse de la saisonnalité des durées entre les événements prix

Comme le laisse supposer l'estimation de l'intensité des événements-prix, les durées entre ceux-ci doivent se caractériser par une forte saisonnalité. La rareté des événements prix illustre des durées longues tandis qu'une fréquence élevée traduit des durées courtes.

Engle et Russell (1998) supposent que les durées observées peuvent être décomposées en deux facteurs. Un premier facteur déterministe qui correspond au caractère saisonnier que

présentent des variables financières (rentabilité, volatilité...) et un facteur dit « aléatoire » que nous allons tenter d'expliquer par la suite.

Parmi les différentes techniques d'interpolation et d'approximation, la fonction *spline* cubique est recommandée par la littérature moderne³¹. Cette fonction possède la forme qui suit :

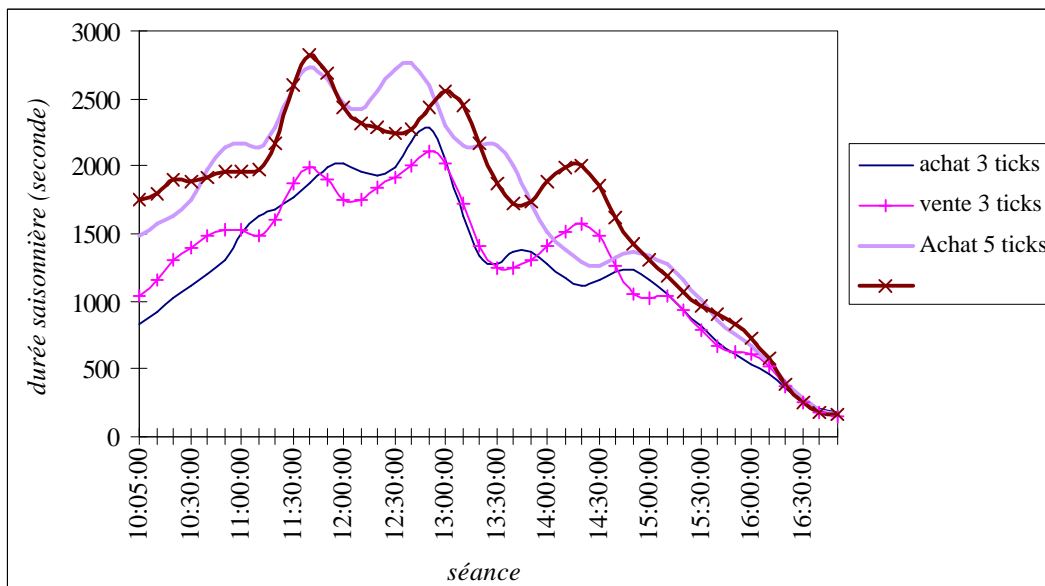
$$f(t_i) = c_j + d_{1,j} * (t_i - x_j) + d_{2,j} * (t_i - x_j)^2 + d_{3,j} * (t_i - x_j)^3$$

où $t_i \in [x_j; x_{j+1}]$ et x_j est le $j^{\text{ème}}$ nœud de la fonction *spline*. Il correspond à la $j^{\text{ème}}$ moyenne de la durée sur l'intervalle de temps considéré.

Dans notre cas, nous avons considéré un intervalle de temps de 15 minutes en moyenne mobile de 5 minutes de manière à éviter certaines discontinuités. Une telle méthode présente néanmoins l'inconvénient de mettre en relief certaines petites discontinuités qui seraient non significatives si le nombre de données dans l'échantillon était plus important³². L'estimation de la composante déterministe des durées entre les événements-prix est illustrée par les graphiques 2.3 et 2.4.

GRAPHIQUE 2.3

Estimation de la composante déterministe de la durée prix de l'Action Pernod-Ricard

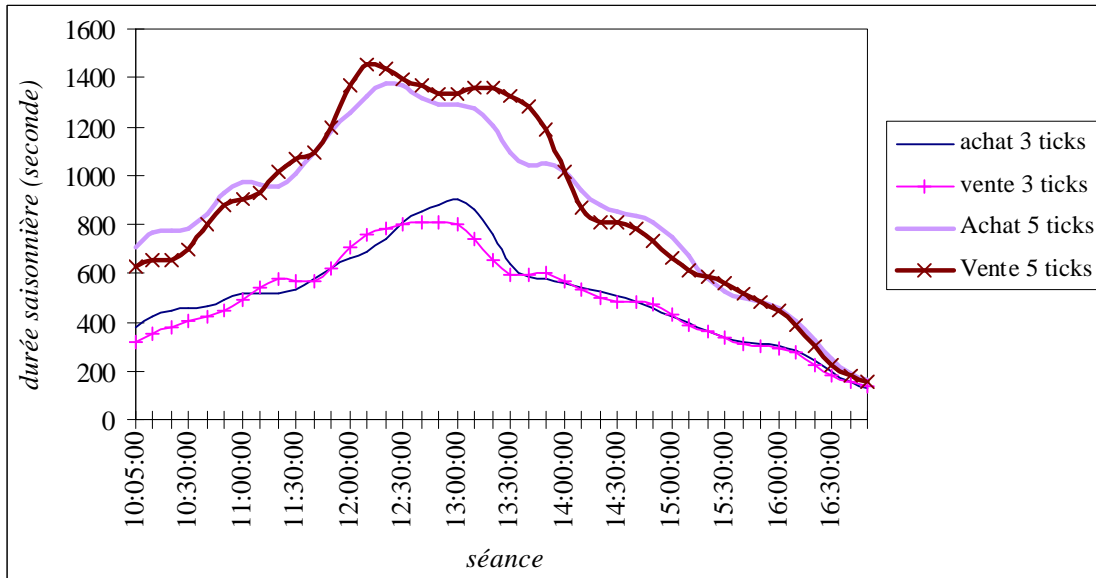


³¹ Voir Engle et Russell (1998).

³² La base employée présente tout de même plus de 100 000 séquences.

GRAPHIQUE 2.4

Estimation de la composante déterministe de la durée prix de l'action Saint-Gobain



L'observation de ces graphiques nous permet de valider une composante saisonnière de tendance semblable quelle que soit l'action considérée même si le profil d'évolution de l'action Pernod-Ricard semble un peu plus irrégulière.

L'estimation de cette composante à chaque instant de la séance, à l'aide de la fonction *spline*, va ainsi nous permettre de calculer les durées désaisonnalisées par une simple division.

Une fois ce travail effectué, il devient nécessaire d'analyser si l'autocorrélation a disparu ou non.

II.2.2 Analyse de l'autocorrélation des durées entre les événements prix

Nous analysons la persistance de l'autocorrélation des durées en dépit de la désaisonnalisation.

Pour cela, nous utilisons la statistique de Ljung-Box. Elle est donnée par la formule suivante :

$$Q_{LB} = N(N + 2) \sum_{j=1}^p \frac{r_j^2}{N - j}$$

où r_j indique la $j^{\text{ème}}$ autocorrélation et N représente le nombre de d'observations. P est le nombre d'autocorrélations. La statistique de Ljung-Box teste l'hypothèse selon laquelle toutes les autocorrélations sont nulles, i.e. la série est de bruit blanc. Sous cette hypothèse nulle, Q est distribuée selon un nombre de degrés de liberté égal au nombre d'autocorrélations³³.

Le tableau 2.2 suivant répertorie la statistique de Ljung-Box pour 10 décalages³⁴. Nous pouvons nous apercevoir que l'autocorrélation diminue mais persiste tout de même. Ce résultat n'est cependant pas valable pour les variations de trois *ticks* de l'action Pernod-Ricard quel que soit le sens de la transaction ; en effet, l'autocorrélation augmente après la désaisonnalisation. Nous rencontrons là un inconvénient majeur connu de la désaisonnalisation.

TABLEAU 2.2

Analyse de l'autocorrélation des durées prix

Action	Seuil	Sens	Q(10)	Q(10) désaisonnalisé
Pernod-Ricard	3 <i>ticks</i>	Achat	263,02	448,09
Pernod-Ricard	3 <i>ticks</i>	Vente	259,53	291,64
Pernod-Ricard	5 <i>ticks</i>	Achat	111,85	89,46
Pernod-Ricard	5 <i>ticks</i>	Vente	75,71	70,698
Saint-Gobain	3 <i>ticks</i>	Achat	404,30	262,46
Saint-Gobain	3 <i>ticks</i>	Vente	619,05	315,93
Saint-Gobain	5 <i>ticks</i>	Achat	191,98	153,58
Saint-Gobain	5 <i>ticks</i>	Vente	209,18	137,30

Un résultat non indiqué dans le tableau 2.2 est le signe des autocorrélations pour chaque décalage. Celui-ci est, à quelques exceptions près, en général positif. Ceci est une preuve supplémentaire de l'existence de phase d'intense activité et de calme.

Devant la persistance du caractère autorégressif des « durée-prix », nous devons rechercher à présent une modélisation adéquate afin de capter l'autocorrélation restante. Pour cela, on va avoir recours aux modèles de durée de type ACD.

³³ Voir aussi Box and Pierce.

³⁴ Il n'existe pas de règle en matière de choix du nombre de décalages. Nous adoptons une attitude semblable à celle de Engle et Russell (1998) et Bauwens et Giot (1999).

II.3 DÉTERMINATION DU MODÈLE DE DURÉE

Nous allons tenter de modéliser la durée qui sépare deux évènements prix consécutifs. Pour cela, nous allons mettre en œuvre quatre modèles présentés dans la section 1: les modèles EACD et WACD ainsi que la première³⁵ des deux versions du modèle log-ACD proposées par Bauwens et Giot (1999). Nous ne reportons que les résultats concernant les versions avec un seul retard dans la mesure où nous avons constaté que des problèmes de convergence et de significativité se posaient pour une très large majorité des autres modèles testés.

Le « meilleur » modèle est celui qui permet de capter au mieux les autocorrélations. Bien entendu, nous allons vérifier que les durées standardisées sont indépendantes, identiquement distribuées et suivent la loi adéquate. Pour cela, nous effectuons un test de non-excès de dispersion, i.e. l'écart-type est égal à la moyenne elle même égale à l'unité. Cette statistique peut être formulée de la manière suivante :

$$\sqrt{N} \left(\frac{\hat{\mathbf{s}}_e^2 - 1}{\mathbf{s}_v} \right)$$

où $\hat{\mathbf{s}}_e^2$ est la variance des durées standardisées

et \mathbf{s}_v est l'écart type de $(e_i - 1)^2$

Le modèle ainsi retenu sert alors de base à un modèle final prenant en compte d'autres variables endogènes.

II.3.1 Le modèle de durée exponentielle

Nous testons ici le modèle EACD(1,1) qui suit :

$$\mathbf{y}_i = \mathbf{w} + \mathbf{a} * \tilde{X}_{i-1} + \mathbf{b} * \mathbf{y}_{i-1}$$

La fonction Log Vraisemblance est alors la suivante :

³⁵ La modélisation du modèle Log-ACD₂ n'a pas été rendue possible pour certains flux en raison de la mauvaise estimation de la log-vraisemblance.

$$L(X_1, \dots, X_{N(t)}) = \sum_{i=1}^{N(t)} -\ln(y_i) - \left(\frac{X_i}{y_i} \right)$$

Les résultats de nos estimations sont reportés dans les tableaux 2.3 et 2.4 qui suivent.

TABLEAU 2.3

Modélisation EACD(1,1) des « durées-prix » de l'action Pernod-Ricard

Sens	Achat	Vente	Achat	Vente
Seuil	3 ticks	3 ticks	5 ticks	5 ticks
<i>w</i>	0,042 (4,263)	0,120 (4,168)	0,115 (2,904)	0,183 (3,182)
a	0,211 (8,220)	0,190 (6,694)	0,184 (5,325)	0,153 (4,584)
b	0,759 (27,92)	0,701 (14,92)	0,713 (11,63)	0,681 (9,414)
Moyenne	1,019	1,000	1,001	1,003
Écart type	1,235	1,236	1,288	1,305
Ljung Box	15,524	4,927	13,368	9,884
Statistique de dispersion	3,240*	3,563*	2,950*	2,681*

Entre parenthèses figure la statistique de Student
* : Non significatif au seuil de 10%

TABLEAU 2.4

Modélisation EACD(1,1) des « durées-prix » de l'action Saint-Gobain

Sens	Achat	Vente	Achat	Vente
Seuil	3 ticks	3 ticks	5 ticks	5 ticks
<i>w</i>	0,132 (3,599)	0,148 (5,701)	0,177 (2,742)	0,268 (3,972)
a	0,133 (6,311)	0,166 (9,187)	0,161 (4,668)	0,216 (6,760)
b	0,742 (14,15)	0,694 (18,49)	0,672 (7,388)	0,538 (6,325)
Moyenne	0,999	1,000	0,999	1,000
Écart type	1,158	1,164	1,197	1,236
Ljung Box	12,874	9,142	14,615	20,088
Statistique de dispersion	3,021*	4,055*	3,061*	4,071*

Entre parenthèses figure la statistique de Student
* : Non significatif au seuil de 10%

Quels que soient les flux considérés, tous les coefficients des modèles sont significatifs au seuil de 1%. De plus, nous pouvons remarquer que ces paramètres sont positifs ce qui est une preuve supplémentaire de la concentration des évènements-prix: une durée courte,

respectivement longue, entre deux variations de prix significatives entraîne une durée courte, respectivement longue, entre les deux évènements-prix suivants.

La somme des coefficients alpha et bêta est inférieure à l'unité, ce qui permet d'assurer l'existence de la moyenne inconditionnelle [voir Engle et Russell (1998)].

Le modèle semble capturer l'autocorrélation des durées de manière efficace dans la mesure où nous acceptons l'hypothèse d'absence d'autocorrélation au seuil de 10% pour l'ensemble des estimations. Le statistique de Ljung Box est en effet inférieure à sa valeur critique.

La validité du modèle EACD(1,1) construit repose sur l'hypothèse de la loi suivie par les durées standardisées. Ces dernières doivent suivre une loi exponentielle. Or le test de non-excès de dispersion nous conduit à rejeter l'hypothèse de dispersion nulle au niveau de confiance de 10%. Les durées standardisées ne suivent donc pas une loi exponentielle. Par conséquent, nous devons rejeter la modélisation des évènements-prix par un modèle EACD(1,1).

II.3.2 Le modèle de durée de type weibull

Nous allons tester le modèle WACD(1,1) suivant :

$$y_i = w + a * \tilde{X}_{i-1} + b * y_{i-1}$$

La fonction Log Vraisemblance est alors la suivante :

$$L(X_1, \dots, X_{N(t)}) = \sum_{i=1}^{N(t)} \ln\left(\frac{g}{X_i}\right) + g * \ln\left(\frac{\Gamma(1+1/g) * X_i}{y_i}\right) - \left(\frac{\Gamma(1+1/g) * X_i}{y_i}\right)^g$$

Les résultats de nos estimations sont reportés dans les tableaux 2.5 et 2.6.

A l'instar des résultats obtenus par les modèles EACD(1,1), les coefficients sont significativement positifs. De même, l'autocorrélation a sensiblement été réduite puisque les statistiques de Ljung Box sont en deçà de leur valeur critique.

Le coefficient gamma est proche de l'unité, mais il en est significativement inférieur. Les modèles WACD(1,1) sont donc significativement différents des modèles EACD(1,1).

En revanche, contrairement aux premiers modèles testés, les durées standardisées suivent une loi de weibull, i.e. les durées transformées suivent une loi exponentielle. Ce résultat est

valable pour l'ensemble des flux de transactions testés. Ceci signifie que l'hypothèse de non-excès de dispersion étant acceptée, nous pouvons valider les modélisations WACD(1,1).

TABLEAU 2.5

Modélisation WACD(1,1) des « durées-prix » de l'action Pernod-Ricard

Sens	Achat	Vente	Achat	Vente
Seuil	3 ticks	3 ticks	5 ticks	5 ticks
<i>w</i>	0,041 (3,677)	0,120 (3,575)	0,129 (2,458)	0,199 (2,598)
a	0,223 (7,355)	0,202 (5,819)	0,213 (4,538)	0,170 (4,055)
b	0,751 (24,04)	0,691 (12,43)	0,671 (8,117)	0,649 (6,652)
g	0,858 (51,52)	0,848 (49,14)	0,832 (43,11)	0,829 (41,12)
Moyenne	0,936	0,930	0,922	0,922
Écart type	0,972	0,964	0,969	0,974
Ljung Box	14,387	4,540	11,914	9,291
Statistique de dispersion	-0,688	-0,971	-0,609	-0,476

Entre parenthèses figure la statistique de Student

TABLEAU 2.6

Modélisation WACD(1,1) des « durées-prix » de l'action Saint-Gobain

Sens	Achat	Vente	Achat	Vente
Seuil	3 ticks	3 ticks	5 ticks	5 ticks
<i>w</i>	0,137 (3,455)	0,148 (5,345)	0,184 (2,796)	0,284 (4,178)
a	0,138 (5,991)	0,169 (8,600)	0,169 (4,691)	0,235 (6,727)
b	0,733 (12,92)	0,691 (17,14)	0,656 (7,090)	0,505 (5,912)
g	0,916 (83,16)	0,925 (86,47)	0,912 (64,01)	0,880 (64,98)
Moyenne	0,962	0,967	0,960	0,946
Écart type	1,008	1,030	1,032	1,012
Ljung Box	12,021	8,302	14,279	19,016
Statistique de dispersion	0,223	1,025	0,718	0,334

Entre parenthèses figure la statistique de Student

II.3.3 Le modèle logarithmique de durée

Nous testons à présent la première version Log-ACD(1,1) proposée par Bauwens et Giot (1998).

La fonction de log vraisemblance est identique à celle du modèle WACD. Il s'agit de la fonction suivante :

$$L = \sum_{i=1}^{N(t)} \ln(\mathbf{g}) - \ln(X_i) + \mathbf{g} \ln[X_i \cdot \Gamma(1+1/\mathbf{g})] - \mathbf{g}\Psi_i - \left[\frac{X_i \cdot \Gamma(1+1/\mathbf{g})}{e^{\Psi_i}} \right]^{\mathbf{g}}$$

Nous considérons ainsi le modèle Log-ACD₁ (1,1) suivant :

$$\Psi_i = \mathbf{w}' + \mathbf{a}Ln(\mathbf{e}_{i-1}) + (\mathbf{a} + \mathbf{b})\Psi_{i-1} \quad \text{Où } \mathbf{w}' = \mathbf{w} + \mathbf{a}Ln[\Gamma(1+1/\mathbf{g})]$$

Ou encore $\Psi_i = \mathbf{w} + \mathbf{a} \cdot Ln(X_{i-1}) + \mathbf{b} \cdot \Psi_{i-1}$

Les résultats de nos estimations sont reportés dans les tableaux 2.7 et 2.8.

Ce dernier modèle testé aboutit lui aussi à des résultats positifs puisque les coefficients sont significativement positifs, l'autocorrélation des durées est capturée. Enfin, les durées transformées suivent une loi exponentielle, ce qui signifie que les durées standardisées entre évènements-prix consécutifs suivent bien une loi de weibull. La modélisation LOG-ACD₁ (1,1) est donc elle aussi validée.

TABLEAU 2.7

Modélisation LOG-ACD₁(1,1) des « durées-prix » de l'action Pernod-Ricard

Sens	Achat	Vente	Achat	Vente
Seuil	3 ticks	3 ticks	5 ticks	5 ticks
<i>w</i>	0,090 (7,287)	0,106 (6,244)	0,099 (4,692)	0,135 (4,538)
<i>a</i>	0,132 (8,460)	0,147 (7,570)	0,136 (5,368)	0,162 (6,028)
<i>b</i>	0,789 (27,06)	0,699 (13,83)	0,711 (9,442)	0,500 (4,402)
<i>g</i>	0,861 (51,47)	0,843 (49,35)	0,838 (43,07)	0,833 (41,26)
Moyenne	0,938	0,927	0,926	0,924
Écart type	0,976	0,972	0,977	0,982
Ljung Box	14,609	13,119	13,733	11,729
Statistique de dispersion	-0,569	-0,699	-0,399	-0,335

Entre parenthèses figure la statistique de Student

TABLEAU 2.8

Modélisation LOG-ACD₁(1,1) des « durées-prix » de l'action Saint-Gobain

Sens	Achat	Vente	Achat	Vente
Seuil	<i>3 ticks</i>	<i>3 ticks</i>	<i>5 ticks</i>	<i>5 ticks</i>
<i>w</i>	0,072 (6,915)	0,074 (8,186)	0,098 (5,978)	0,108 (7,324)
a	0,107 (7,886)	0,119 (9,721)	0,150 (7,962)	0,157 (9,401)
b	0,714 (13,38)	0,701 (16,86)	0,517 (5,862)	0,594 (10,32)
g	0,917 (83,19)	0,921 (86,86)	0,914 (64,054)	0,890 (64,96)
Moyenne	0,963	0,965	0,961	0,951
Écart type	1,006	1,045	1,033	1,013
Ljung Box	13,888	17,184	15,991	22,829
Statistique de dispersion	0,181	1,404	0,799	0,371

Entre parenthèses figure la statistique de Student

Il convient maintenant de comparer ces résultats avec ceux obtenus à l'aide des modèles WACD(1,1). Nous constatons, à l'aide de la statistique de Ljung-Box, que les modèles WACD(1,1) semblent mieux capturer l'autocorrélation des durées. Ces résultats sont conformes à ceux obtenus par Bauwens et Giot(1999). En revanche, Aubier (2000) observe des différences non significatives entre les modèles³⁶.

Nous pouvons dès lors regretter la non-convergence du modèle LOG-ACD₂(1,1) puisque ce modèle semble être le plus performant selon Bauwens et Giot (1999). Ceci est d'autant plus regrettable que le modèle WACD(1,1) est soumis à des contraintes de signes plus restrictives, ce qui peut engendrer des problèmes de convergence lors de l'intégration de nouvelles variables explicatives.

La statistique VNET va donc être construite à l'aide des espérances de durées estimées à partir d'un modèle WACD(1,1) que nous allons tenter d'améliorer au préalable.

³⁶ Il convient de signaler que l'analyse d'Aubier (2000) ne porte que sur une action, i.e. Saint-Gobain, durant une période d'étude très courte. De plus, l'auteur définit l'événement-prix comme étant une variation de deux échelons de cotation du milieu de la fourchette, ce qui peut d'ailleurs apparaître comme très faible.

II.4 EXTENSION DU MODÈLE DE DURÉE

Nous avons mis en relief la capacité du modèle WACD(1,1) à capturer l'autocorrélation des durées standardisées entre les évènements prix. Il faut maintenant essayer d'améliorer ce modèle en ajoutant de nouvelles variables dans l'équation à l'instar de Engle et Russell (1998) ou encore Bauwens et Giot (1999).

II.4.1 Propositions de variables additionnelles

Nous allons ainsi introduire dans la modélisation WACD. La présence de cinq variables est ainsi testée. Ces variables³⁷ sont les suivantes:

◆ **Le volume anormal**, noté *VOLA*, désigne le volume standardisé. L'idée sous-jacente est qu'un accroissement imprévu des volumes sur le marché signale la présence d'arrivée d'informations, ce qui n'est bien entendu pas sans conséquence sur le processus de formation des prix. Ainsi, nous pouvons penser que le lien entre la « durée-prix » et le volume anormal est négatif puisque l'arrivée d'informations devrait provoquer un accroissement de la volatilité et du nombre de cotations.

◆ **La fréquence des appariements**, notée *FNB*, désigne le nombre de transactions par seconde. Comme il l'a été signalé lors de la section théorique, on peut penser que la volatilité est plus élevée après des périodes de forts taux de transaction. Nous nous situons ainsi dans le cadre théorique formulé par Easley et O'Hara (1992). Une relation négative devrait ainsi être obtenue entre l'espérance de « la durée-prix » et l'intensité des transactions.

◆ **Le déséquilibre entre les volumes des transactions acheteuses et vendeuses dans l'intervalle de deux évènements-prix**, noté *DSQ*, a été proposé par Bauwens et Giot (1999). Il illustre l'existence d'une tendance dans la formation des prix. On a estimé cette mesure de manière identique à celle proposée par ses inspirateurs.

Le déséquilibre durant la $i^{\text{ème}}$ « durée-prix » se formule ainsi de la manière qui suit :

³⁷ En fait, nous avons testé d'autres variables, i.e. la taille moyenne des transactions, le déséquilibre entre les volumes associés aux meilleures limites, l'intensité anormale des appariements. Nos résultats n'étant absolument pas contributifs à une meilleure prévision en raison de leur non significativité, de la non-convergence du modèle ou encore de l'étroite corrélation avec l'une des cinq variables testées, nous avons choisi de ne pas les présenter.

$$DSQ_i = \frac{\sum_j A_j Vol_j - \sum_j V_j Vol_j}{\sum_j Vol_j}$$

où A_j et V_j sont des variables indicatrices. A_j vaut 1 si la transaction est un achat et 0 sinon, inversement, V_j vaut 1 si la transaction est une vente et 0 sinon. Et Vol_j est le volume échangé lors de la $j^{\text{ème}}$ transaction au sein de la « durée-prix » i .

La relation que l'on va obtenir devrait être négative ; en effet, nous pouvons penser que l'existence d'une pression acheteuse ou vendeuse signale des mouvements de prix futur d'où des « durées-prix » rapprochées.

♦ **La fourchette nominale**, notée F_NOM , constitue une mesure de la liquidité. De plus, une fourchette large peut signifier une incertitude plus importante sur la valeur fondamentale de l'action. Par conséquent, à double titre, nous pouvons supposer qu'il existe une relation négative entre la taille de la fourchette et la « durée-prix ».

♦ **Le lambda de Kyle**, noté LK , est une autre mesure de la liquidité que nous avons décrite dans le premier chapitre. De la même manière que pour la fourchette, on émet l'hypothèse d'une relation négative. Dans un marché peu liquide, les variations de prix sont plus nombreuses ; ainsi, les durées qui séparent les événements-prix vont être plus courtes.

On va donc estimer l'espérance conditionnelle de la « durée Prix » pour chaque flux à l'aide d'un modèle WACD(1,1) auquel on rajoute le niveau d'une des variables décrites ci-dessus, décalé d'un retard. Ce modèle est formulé ci-après.

$$y_i = w + a * \tilde{X}_{i-1} + b * y_{i-1} + dVar_{i-1} \quad (2.14)$$

Avec \tilde{X}_i , la durée désaisonnalisée qui correspond au temps séparant le $i^{\text{ème}}$ et le $(i-1)^{\text{ème}}$ événement prix.

où y_i n'est autre que l'espérance conditionnelle de la durée sachant son passé.

et Var_i correspond au niveau de la variable de microstructure retenue lors de l' $i^{\text{ème}}$ événement-prix.

La fonction Log Vraisemblance reste identique au modèle de base :

$$L(X_1, \dots, X_{N(t)}) = \sum_{i=1}^{N(t)} \text{Ln} \left(\frac{g}{X_i} \right) + g^* \text{Ln} \left(\frac{\Gamma(1+1/g)^* X_i}{y_i} \right) - \left(\frac{\Gamma(1+1/g)^* X_i}{y_i} \right)^g$$

II.4.2 Résultats et Interprétations

Les résultats de nos modélisations sont répertoriés dans les tableaux 2.9 et 2.10 qui suivent.

Le premier commentaire général concerne le sens de la relation qui lie ces nouvelles variables explicatives à l'espérance de la durée entre deux évènements-prix. Pour chacune de ces variables, lorsque le coefficient est significatif, la corrélation est négative. Ceci valide donc les hypothèses effectuées précédemment. Un élargissement de la fourchette a pour conséquence la réduction de la « durée-prix ». Le lambda de Kyle, estimateur de la liquidité, indique que lorsque la liquidité se dégrade, les variations de prix significatives se succèdent plus rapidement. La présence d'un déséquilibre dans le sens des volumes échangés indique une tendance prononcée du processus de formation des prix, et donc un accroissement des évènements-prix, l'intensité des transactions, bien qu'elle puisse exprimer une augmentation de la liquidité, amplifie les mouvements de prix d'où le lien négatif; ceci corrobore la conjecture émise par Easley et O'Hara (1992) selon laquelle l'absence de transaction signifie qu'aucune nouvelle arrive sur le marché. Enfin, un volume anormalement élevé est le signe de la présence d'investisseur informé d'où des mouvements de prix rapprochés.

En ce qui concerne la réduction de l'autocorrélation, nous devons constater que celle-ci diminue très faiblement quand elle ne remonte pas ! L'intégration de la fourchette nominale, du volume anormal et de l'intensité des échanges semble apporter le plus de satisfactions.

Si nous observons au cas par cas ces résultats, nous nous apercevons que certaines variables sont régulièrement plus significatives que d'autres. Ceci est particulièrement vrai pour le volume anormal, l'intensité des appariements et la fourchette nominale. En ce qui concerne le lambda de Kyle et le déséquilibre mesuré à la Bauwens et Giot (1999), les coefficients sont souvent non significatifs. Enfin, nous pouvons remarquer que ces problèmes de significativité sont plus récurrents pour l'action Pernod-Ricard.

TABLEAU 2.9

Modélisation WACD(1,1) des durées entre variations de cours de l'action Pernod-Ricard avec introduction de variables explicatives supplémentaires

Flux	Achat	Achat	Vente	Vente
Seuil	3 ticks	5 ticks	3 ticks	5 ticks
Fourchette Nominale				
<i>w</i>	0,075 (4,296)	0,204 (3,871)	0,101 (2,742)	0,177 (2,054)
<i>a</i>	0,209 (7,323)	0,185 (4,705)	0,205 (5,649)	0,170 (3,964)
<i>b</i>	0,752 (24,97)	0,687 (10,76)	0,686 (11,84)	0,654 (6,610)
<i>g</i>	0,861 (51,32)	0,841 (42,83)	0,847 (49,15)	0,829 (41,10)
<i>d</i>	-0,071 (-3,44)	-0,181 (-4,621)	0,052 (0,969)	0,041 (0,488)
Q(10)	13,124	10,424	ns	ns
Lambda de Kyle				
<i>w</i>	0,045 (3,420)	0,132 (2,899)	0,115 (3,221)	0,275 (1,686)
<i>a</i>	0,223 (7,436)	0,205 (4,856)	0,200 (5,747)	0,189 (3,531)
<i>b</i>	0,749 (24,44)	0,692 (9,871)	0,693 (12,44)	0,579 (3,417)
<i>g</i>	0,858 (51,46)	0,832 (43,08)	0,847 (49,13)	0,828 (41,14)
<i>d</i>	-0,167 (-0,766)	-0,566 (-0,914)	0,118 (0,313)	-0,686 (-0,782)
Q(10)	ns	ns	ns	ns
Déséquilibre				
<i>w</i>	0,061 (1,997)	0,185 (2,401)	0,025 (0,577)	0,139 (1,604)
<i>a</i>	0,219 (7,276)	0,208 (4,529)	0,220 (53,99)	0,174 (4,038)
<i>b</i>	0,753 (24,42)	0,674 (8,708)	0,673 (12,46)	0,653 (7,050)
<i>g</i>	0,857 (51,52)	0,832 (43,08)	0,850 (49,03)	0,829 (41,09)
<i>d</i>	-0,029 (-0,726)	-0,086 (-1,071)	-0,144 (-2,429)	0,081 (0,983)
Q(10)	ns	ns	6,174	ns
Intensité des appariements				
<i>w</i>	0,056 (4,515)	n/c	n/c	0,281 (3,907)
<i>a</i>	0,196 (6,374)	n/c	n/c	0,175 (4,332)
<i>b</i>	0,762 (22,99)	n/c	n/c	0,581 (6,801)
<i>g</i>	0,860 (51,48)	n/c	n/c	0,831 (41,16)
<i>d</i>	-0,230 (-4,410)	n/c	n/c	-0,671 (-8,975)
Q(10)	15,381	n/c	n/c	12,125
Volume Anormal				
<i>w</i>	0,061 (4,225)	0,206 (2,920)	0,130 (4,167)	0,192 (2,655)
<i>a</i>	0,219 (7,112)	0,241 (4,635)	0,209 (5,963)	0,170 (3,876)
<i>b</i>	0,744 (22,34)	0,614 (6,268)	0,697 (13,42)	0,683 (7,280)
<i>g</i>	0,861 (51,43)	0,838 (43,13)	0,850 (49,20)	0,831 (41,11)
<i>d</i>	-0,013 (-4,528)	-0,030 (-4,172)	-0,020 (-4,944)	-0,017 (-2,660)
Q(10)	13,326	9,152	4,151	8,190

Entre parenthèses figure la statistique de Student, n/c indique que le modèle ne converge pas tandis que ns signifie que la statistique de Ljung-Box est rejetée.

TABLEAU 2.10

Modélisation WACD(1,1) des durées entre variations de cours de l'action Saint-Gobain avec introduction de variables explicatives supplémentaires

Flux	Achat	Achat	Vente	Vente
Seuil	3 ticks	5 ticks	3 ticks	5 ticks
Fourchette Nominale				
<i>w</i>	0,190 (4,352)	0,319 (5,241)	0,185 (5,438)	0,395 (5,608)
<i>a</i>	0,133 (6,619)	0,175 (6,508)	0,169 (9,017)	0,229 (7,505)
<i>b</i>	0,717 (13,81)	0,572 (8,396)	0,673 (16,72)	0,460 (6,623)
<i>g</i>	0,917 (83,15)	0,914 (64,06)	0,925 (86,43)	0,882 (65,00)
<i>d</i>	-0,075 (-3,612)	-0,113 (-6,408)	-0,046 (-2,431)	-0,126 (-4,940)
Q(10)	10,932	9,888	7,176	17,002
Lambda de Kyle				
<i>w</i>	0,146 (3,695)	0,227 (3,968)	0,191 (7,113)	n/c
<i>a</i>	0,133 (6,043)	0,175 (5,966)	0,176 (9,635)	n/c
<i>b</i>	0,736 (13,43)	0,622 (8,588)	0,658 (18,20)	n/c
<i>g</i>	0,916 (83,13)	0,912 (64,03)	0,926 (86,45)	n/c
<i>d</i>	-0,635 (-1,972)	-0,994 (-2,036)	-1,205 (-7,041)	n/c
Q(10)	12,011	10,214	8,553	n/c
Déséquilibre				
<i>w</i>	0,216 (3,983)	0,230 (2,934)	0,137 (3,668)	0,339 (4,383)
<i>a</i>	0,131 (6,053)	0,162 (4,628)	0,170 (8,506)	0,207 (5,570)
<i>b</i>	0,720 (12,84)	0,652 (7,204)	0,691 (17,16)	0,541 (6,232)
<i>g</i>	0,917 (83,09)	0,912 (64,01)	0,924 (86,46)	0,881 (64,94)
<i>d</i>	-0,100 (-2,629)	-0,065 (-1,270)	0,014 (0,399)	-0,121 (-1,893)
Q(10)	11,133	ns	ns	ns
Intensité des appariements				
<i>w</i>	0,195 (5,558)	0,210 (3,806)	0,189 (7,202)	n/c
<i>a</i>	0,119 (6,661)	0,130 (4,470)	0,146 (8,269)	n/c
<i>b</i>	0,715 (16,38)	0,691 (9,327)	0,691 (19,32)	n/c
<i>g</i>	0,919 (83,22)	0,916 (63,94)	0,929 (86,62)	n/c
<i>d</i>	-0,415 (-6,487)	-0,517 (-8,159)	-0,365 (-14,41)	n/c
Q(10)	17,546	7,124	6,587	n/c
Volume Anormal				
<i>w</i>	0,183 (5,643)	0,283 (4,083)	0,190 (6,927)	0,355 (5,961)
<i>a</i>	0,162 (7,879)	0,230 (6,203)	0,203 (9,932)	0,304 (7,684)
<i>b</i>	0,700 (16,19)	0,547 (6,471)	0,651 (17,58)	0,436 (6,079)
<i>g</i>	0,919 (83,25)	0,914 (63,97)	0,928 (86,45)	0,883 (65,01)
<i>d</i>	-0,088 (-6,390)	-0,063 (-3,326)	-0,089 (-6,269)	-0,094 (-5,294)
Q(10)	12,598	13,115	10,456	18,995

Entre parenthèses figure la statistique de Student, n/c indique que le modèle ne converge pas tandis que ns signifie que la statistique de Ljung-Box est rejetée.

Ces premiers résultats obtenus, nous avons souhaité intégrer les variables significatives individuellement au sein d'une même modélisation WACD(1,1) afin de réduire au maximum l'autocorrélation des durées standardisées évaluée à l'aide de la statistique de Ljung-Box, et améliorer la prévision de l'espérance conditionnelle de la « durée prix ». Malheureusement, nous avons fait face à des problèmes de convergence dans la mesure où le choix du modèle WACD(1,1), le meilleur dans le cas présent, implique des contraintes de positivité des signes.

Nous ne pouvons donc retenir que des modèles bénéficiant de l'introduction d'une seule variable explicative supplémentaire. Pour des raisons de lisibilité, nous avons retenu la même modélisation par action afin d'estimer la statistique VNET. Ainsi, pour l'action Pernod-Ricard, nous avons choisi d'intégrer le volume anormal tandis que pour l'action Saint-Gobain, c'est la fourchette nominale qui a été choisie.

II.5 MODÉLISATION DE LA VNET

L'objectif poursuivi est d'appliquer au marché français la statistique VNET proposée par Engle et Lange (1998). Comme les auteurs, nous testons la présence de plusieurs variables explicatives.

L'espérance conditionnelle de la durée est estimée à partir du modèle WACD obtenu précédemment.

II.5.1 Présentations des régressions

Engle et Lange (1997) analysent la profondeur du marché à travers une statistique nommée VNET. Cette statistique désigne l'excès de demande qui peut être échangé sans provoquer de variation de prix jugée significative. Nous reprenons une formalisation identique à celle des auteurs :

$$VNET = Ln \left| \sum_i (d_i * vol_i) \right|$$

où d_i correspond à la variable indicatrice représentant la direction de la transaction. Si l'échange est effectué sur l'initiative d'un acheteur (vendeur), alors $d_i = +1$ ($d_i = -1$), et Vol_i indique le volume échangé lors de la transaction i pendant une « durée-prix » donnée.

Une fois construite cette statistique, notre objectif consiste à expliquer son évolution durant la séance de cotation par des variables de microstructure. Comme l'indique le tableau 2.11, nous nous apercevons qu'une large majorité de l'excès de demande est inférieure à 1000 titres lorsque le seuil définissant l'événement prix est de trois *ticks*. Quand celui-ci est de 5 échelons de cotation, nous apercevons que la médiane est inférieure à 2000 actions.

TABLEAU 2.11

Statistiques Descriptives concernant le niveau de la statistique VNET

Action	Flux	Seuil	Moyenne	Médiane	Écart-type
Saint-Gobain	Achat	3 ticks	2529	1220	3720
Saint-Gobain	Vente	3 ticks	2393	1101	3786
Pernod-Ricard	Achat	3 ticks	3060	1233	5263
Pernod-Ricard	Vente	3 ticks	3257	1293	6176
Saint-Gobain	Achat	5 ticks	3652	1892	4976
Saint-Gobain	Vente	5 ticks	3596	1850	5086
Pernod-Ricard	Achat	5 ticks	3960	1646	7065
Pernod-Ricard	Vente	5 ticks	4243	1715	7578

A l'instar d'Engle et Lange (1997), on va mettre en œuvre des régressions linéaires. Nous allons intégrer la série de variables suivantes³⁸ dans leur modèle explicatif :

- ◆ **La taille de la fourchette nominale lors de l'événement-prix précédent**, notée F_NOM au titre de l'asymétrie d'information.
- ◆ **Le lambda de Kyle à la fin de la « durée-prix » précédente**, noté LK . Il s'agit d'une autre mesure de la liquidité ou encore de l'asymétrie d'information.
- ◆ **Le déséquilibre entre les volumes des transactions associés aux meilleures limites acheteuses et vendeuses lors de la fin de la « durée-prix » précédente**, noté $VNCO$.

³⁸ Nous considérons les transformées logarithmiques de ces variables suivant ainsi la méthode des instigateurs de ce modèle de prévision.

Cette variable peut être considérée comme un signe d'excès de demande future. Cette mesure peut être formulée de la manière suivante :

$$VNCO = Ln|(C_i * vol_i)|$$

Où C_i désigne la variable indicatrice représentant la borne à laquelle est associée le volume Vol_i . S'il s'agit de la meilleure limite acheteuse (vendeuse), alors $C_i = +1$ ($C_i = -1$).

◆ **L'intensité des transactions pendant la « durée-prix » précédente**, notée *FNB*, qu'il convient de ne pas confondre avec la variable nombre de transactions. On justifie ce choix par l'interprétation ambiguë qui peut être faite de cette dernière variable. En effet, un nombre de transactions important dans une durée courte n'illustre pas la même réalité qu'un nombre élevé de transactions au sein d'une durée longue. Cette variable paraît mieux refléter l'hypothèse de la présence des investisseurs informés sur le marché.

◆ **Le volume de transactions anormal pendant la « durée-prix » précédente**, noté *VOLA*. Cette variable permet de détecter une arrivée d'information sur le marché. Nous savons que le volume anormal réduit la durée entre les événements prix. Son incidence sur le déséquilibre est plus incertain dans la mesure où il n'est pas signé.

◆ **La taille moyenne des transactions durant la « durée-prix » précédente**, notée *TMT*. Cette variable est un signal de la qualité de l'information parvenant au marché.

◆ **L'écart de prix justifiant le nouvel événement-prix**, noté *SAUT*. Cet écart peut être supérieur au seuil définissant « l'événement-prix » pour le cas où la transaction qui déclenche le seuil touche plusieurs limites.

◆ **L'espérance conditionnelle de la durée**, notée *EspT*. Il s'agit de l'espérance conditionnelle de la « durée Prix » pour chaque flux de transaction.

◆ **L'erreur de prévision**, notée *EET*. Elle correspond à la durée standardisée obtenue dans le cadre de la modélisation WACD.

Les variables explicatives étant définies, on peut à présent effectuer les estimations des coefficients de ces régressions.

II.5.2 Résultats des régressions et interprétation

Nous avons tout d'abord effectué plusieurs simulations à partir de chaque flux de transactions selon une variation de prix fixée à l'avance - trois ou cinq échelons de cotations. Ces résultats sont répertoriés dans le tableau 2.12.

Nous pouvons nous apercevoir que de nombreux coefficients ne sont pas significatifs ; ceci est particulièrement le cas pour les variables de liquidité que sont la fourchette nominale et le lambda de Kyle. Cela est aussi vrai pour la variable représentant le déséquilibre des volumes au sein du carnet d'ordres ainsi que l'espérance conditionnelle de la « durée-prix ».

TABLEAU 2.12
Régressions individuelles de la VNET

	<i>Cste</i>	<i>F_NOM</i>	<i>LK</i>	<i>FNB</i>	<i>VNCO</i>	<i>VOLA</i>	<i>TMT</i>	<i>SAUT</i>	<i>EspT</i>	<i>EET</i>
Pernod-Ricard Achat 3 ticks	5,244 (15,78)*	0,364 (1,42)	-0,153 (-1,45)	2,757 (4,25)*	0,031 (0,90)	0,078 (3,40)*	0,222 (5,25)*	0,848 (2,62)*	0,026 (1,16)	0,487 (14,23)*
Pernod-Ricard Vente 3 ticks	5,583 (15,57)*	-0,003 (-0,01)	-0,169 (-1,54)	2,278 (3,28)*	0,028 (0,78)	0,063 (2,84)*	0,198 (4,37)*	0,683 (2,16)*	0,146 (1,48)	0,483 (13,98)*
Pernod-Ricard Achat 5 ticks	5,123 (12,12)*	0,240 (0,86)	-0,027 (-0,22)	4,016 (4,24)*	0,082 (2,06)*	0,031 (1,61)	0,267 (5,27)*	0,534 (1,45)	0,019 (1,04)	0,574 (14,35)*
Pernod-Ricard Vente 5 ticks	6,277 (13,61)*	-0,024 (-0,08)	-0,115 (-0,98)	2,809 (3,18)*	0,040 (0,99)	0,062 (3,16)*	0,143 (2,67)*	0,670 (1,90)	0,097 (1,01)	0,586 (14,32)*
Saint-Gobain Achat 3 ticks	5,452 (23,96)*	0,092 (0,76)	-0,160 (-1,47)	1,434 (5,43)*	0,032 (1,60)	0,215 (5,21)*	0,162 (6,12)*	0,649 (4,26)*	0,259 (1,54)	0,447 (23,40)*
Saint-Gobain Vente 3 ticks	5,160 (25,14)*	0,012 (0,10)	-0,143 (-1,35)	1,719 (6,49)*	0,063 (3,31)*	0,235 (5,87)*	0,095 (3,96)*	0,791 (4,87)*	0,848 (5,78)*	0,552 (28,42)*
Saint-Gobain Achat 5 ticks	5,349 (17,04)*	-0,010 (-0,07)	-0,169 (-1,23)	2,625 (5,43)*	0,059 (2,32)*	0,138 (4,43)*	0,228 (5,98)*	0,383 (1,95)	0,285 (1,41)	0,458 (18,13)*
Saint-Gobain Venet 5 ticks	4,995 (17,48)*	0,057 (0,40)	-0,002 (-0,01)	2,396 (6,16)*	0,069 (2,78)*	0,153 (5,19)*	0,182 (5,19)*	0,320 (1,57)	1,103 (6,32)*	0,562 (22,58)*

Entre parenthèses figure le T de Student

* : significatif au seuil de 5%

Face à ces problèmes de significativité, comme Engle et Lange (1997), nous avons empilé les huit flux d'ordres de notre étude. Ceci devrait permettre de détecter des coefficients plus significatifs et donc de déceler les variables qui contribuent à l'explication de la statistique VNET. Les tableau 2.13 indique les résultats de notre régression.

La première remarque concerne la fourchette nominale. Le coefficient de la fourchette nominale est non significatif. La fourchette ne contribue donc pas à l'explication de la statistique VNET. Ce résultat est quelque peu inattendu. Une estimation empilée d'un nombre plus élevé de flux aurait peut être modifié ce résultat.

L'absence du pouvoir explicatif de la fourchette n'est cependant pas préjudiciable au modèle proposé dans la mesure où le lambda de Kyle, qui correspond à une extension de la mesure de la fourchette est corrélé avec la profondeur estimée par la statistique VNET. Cette relation est négative : ceci signifie que lorsque la liquidité se dégrade, i.e. le lambda de Kyle augmente, l'excès de demande diminue.

TABLEAU 2.13

Estimation³⁹ de la VNET à partir des flux d'ordres
des actions Pernod-Ricard et Saint-Gobain

	Coefficients	Écart-type	Statistique t
Constante	5,171	0,101	51,126
<i>LK</i> _{<i>t</i>-1}	-0,162	0,033	-4,919
<i>VNCO</i> _{<i>t</i>-1}	0,058	0,009	6,413
<i>FNB</i> _{<i>t</i>-1}	1,784	0,146	12,219
<i>VOLA</i> _{<i>t</i>-1}	0,104	0,008	12,815
<i>TMT</i> _{<i>t</i>-1}	0,192	0,012	15,773
<i>SAUT</i> _{<i>t</i>}	0,782	0,069	11,263
<i>EET</i> _{<i>t</i>}	0,382	0,064	5,970
<i>EspT</i> _{<i>t</i>}	0,505	0,009	53,909

Pour ce qui est du déséquilibre des volumes au sein du carnet d'ordres, noté *VNCO*, nous avons obtenu une relation positive significative entre cette variable et la statistique VNET. Ceci signifie que le déséquilibre observé au sein du carnet d'ordres n'annonce pas une pression à la vente ou à l'achat et donc un affaiblissement de la liquidité. Ce résultat est contradictoire à nos attentes. Il peut être expliqué par la qualité de cette variable. Tout d'abord, elle ne prend pas en compte des volumes cachés ; de plus, des volumes peu étoffés en première limite peuvent cacher des volumes plus importants aux autres

³⁹ Pour une meilleure lisibilité, seuls les coefficients significativement supérieurs au seuil de 5% ont été répertoriés.

meilleures limites. Une autre explication plausible consiste à admettre qu'un fort déséquilibre implique la présence de volumes importants sur les meilleures limites et donc une profondeur élevée.

Nous obtenons une relation positive entre la VNET et la taille moyenne des transactions passée ; il en est de même avec le volume anormal (non signé) et l'intensité des transactions. Si on se réfère aux hypothèses d'Easley et O'Hara (1992), pour qui l'arrivée de l'information est véhiculée par les volumes, le résultat obtenu va à l'encontre de leurs attentes. Engle et Lange (1998) justifient assez clairement la détection d'un signe positif dans la relation. En effet, un marché très actif est un marché liquide. Par conséquent, plus les quantités de titres échangées sont élevées, meilleure est la profondeur.

Nous obtenons une relation positive avec la variable *SAUT* qui désigne l'écart de prix effectivement réalisé. Ceci est logique. Plus l'écart de prix est élevé, plus le montant échangeable pour atteindre le seuil d'évènement va être important toutes choses égales par ailleurs.

Enfin, en ce qui concerne la prévision de la « durée-prix » et de son erreur d'estimation, nous retrouvons des résultats analogues à ceux d'Engle et Lange (1998). Les relations sont toutes deux positives. Pour le cas de l'espérance conditionnelle de la « durée-prix », la corrélation positive s'explique par le fait que la « durée-prix » peut être interprétée comme la fonction réciproque de la volatilité. Ainsi un marché volatil, i.e. présentant des durées courtes, va se caractériser par une plus faible liquidité.

Pour ce qui est de l'erreur de prévision, une relation positive est constatée. L'exemple suivant démontre l'existence d'un tel lien. Si on considère le cas d'une « durée prix » plus courte que prévue, cela signifie que la statistique VNET va diminuer. En effet, en pareille situation, le marché se montre plus volatil que prévu ce qui a pour conséquence d'assécher la liquidité du marché.

CONCLUSION

« Les marchés financiers se révèlent être de magnifiques terrains d'expérience pour lesquels il existe des myriades de données »⁴⁰. Les modèles de durée font partie de ces modélisations qui ont été empruntées au monde de la « science dure » pour être adaptées en finance. Parmi ces utilisations, les modèles de durée ont donné lieu à la réalisation d'autres modélisations telles que la création de la statistique VNET par Engle et Lange (1997). Cette statistique correspond à la demande de titres en excès sous une contrainte de prix. Elle répond ainsi à la définition de la liquidité puisqu'elle intègre ses trois facettes : le temps, le volume et le prix.

Nous avons adapté cette statistique au marché français à travers l'examen de deux actions, i.e. Pernod-Ricard et Saint-Gobain. L'analyse individuelle de ces titres ne permet pas d'obtenir une explication très significative de la liquidité mesurée par la statistique VNET. Les résultats s'avèrent bien meilleurs lors de l'estimation empilée ce qui laisse augurer de réussites futures. La liquidité est ainsi expliquée par l'état du carnet d'ordres, l'activité boursière, le montant des positions, la patience des investisseurs à échanger et la volatilité. Ces relations ne constituent pas une découverte mais, contrairement à d'autres études, on perçoit une certaine dynamique entre ces différentes variables.

Au cours de cette première partie dédiée à l'analyse de la liquidité sous l'angle de la microstructure, nous avons pu nous apercevoir que le coût de liquidité n'est pas un simple coût fixe à intégrer dans sa gestion de portefeuille. La variation de la liquidité est en partie expliquée par l'évolution de la volatilité. L'investisseur doit faire face simultanément à une augmentation de la volatilité ainsi qu'à une détérioration du niveau de liquidité. En raison de cette variabilité, on ne doit plus considérer la liquidité seulement comme une simple friction de l'échange mais aussi comme un risque à part entière.

⁴⁰ « Les marchés financiers tirent profit de la physique théorique », Le Monde, 1^{er} septembre 2000.