



L'ajustement à seuil des processus cointégrés

Que sait-on des modèles à trois régimes ?

Frédérique Bec, Mélika Ben Salem

DANS **REVUE D'ÉCONOMIE POLITIQUE** 2004/4 Vol. 114 , PAGES 467 À 488
ÉDITIONS **DALLOZ**

ISSN 0373-2630

DOI 10.3917/redp.144.0467

Date de mise en ligne : 01/04/2009

Article disponible en ligne à l'adresse

<https://shs.cairn.info/revue-d-economie-politique-2004-4-page-467?lang=fr>



Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...
Scannez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



Distribution électronique Cairn.info pour Dalloz.

Vous avez l'autorisation de reproduire cet article dans les limites des conditions d'utilisation de Cairn.info ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Détails et conditions sur cairn.info/copyright.

Sauf dispositions légales contraires, les usages numériques à des fins pédagogiques des présentes ressources sont soumises à l'autorisation de l'Éditeur ou, le cas échéant, de l'organisme de gestion collective habilité à cet effet. Il en est ainsi notamment en France avec le CFC qui est l'organisme agréé en la matière.

L'ajustement à seuil des processus cointégrés que sait-on des modèles à trois régimes ?

Frédérique Bec*
Mélika Ben Salem**

Cet article traite du concept de cointégration à seuil tel qu'il a été introduit par Balke et Fomby en 1997 pour les modèles à seuils à trois régimes. L'objectif premier est de proposer un état des lieux de cette littérature grandissante, soulignant non seulement les avancées récentes de la théorie économétrique, mais également ses limites actuelles. Notre propos sera illustré par une application à la structure par terme des taux d'intérêt américains sur données hebdomadaires pour la période 1960-2003.

cointégration à seuil - modèle à correction d'erreur non linéaire - structure par terme des taux d'intérêt

Threshold adjustment of cointegrated processes What do we know about three-regime models?

This paper focuses on the threshold cointegration concept as introduced by Balke and Fomby in 1997 within three-regime threshold models. The main goal is to propose an overview of this ever growing literature, stressing not only the recent advances in econometric theory but also its current limits. This is illustrated by an application to the term structure of interest rates, using weekly US data from 1960 to 2003.

threshold cointegration - non-linear error correction model - term structure of interest rates

Classification JEL: F31, C12, C22

Les auteurs remercient les participants à la conférence « Développements récents de l'économétrie appliquée à la finance », Université Paris X - Nanterre pour leurs commentaires. Nous remercions également Dag Tjøstheim pour ses remarques utiles. Nous restons bien sûr seules responsables des éventuels oublis et erreurs.

* CREST-ENSAE, Timbre J120 - 3, av. Pierre Larousse, 92245 Malakoff Cedex. Email : bec@ensae.fr

** OEP, Université de Marne-la-Vallée. Email : melika.bensalem@univ-mlv.fr

REP 114 (4) juill.-août 2004

La transposition des résultats de l'économétrie des séries temporelles linéaires aux processus non linéaires est souvent tentante mais presque toujours fallacieuse, en particulier lorsque l'on considère des processus susceptibles de n'être pas stationnaires.

Il existe différentes classes de modèles non linéaires, comme par exemple la classe des modèles à changement de régimes markoviens, celle des modèles ARCH, ou encore celle des modèles auto-régressifs à seuils. Nous nous concentrerons ici sur les modèles auto-régressifs à seuils à transition brutale entre les différents régimes. Plus précisément, nous nous intéresserons aux résultats concernant les modèles auto-régressifs à trois régimes caractérisés par une fonction discontinue de transition entre régimes, *i.e.* le modèle univarié SETAR (Self-Exciting Threshold Auto-Regression), et le modèle vectoriel à correction d'erreur avec ajustement discontinu TVECM (Threshold Vector Error Correction Model). Ce choix est motivé par la pertinence de ce type de modèles pour de nombreuses applications macroéconomiques et financières. Ainsi que l'ont souligné Balke et Fomby [1997], en décrivant un processus d'ajustement discontinu à l'équilibre, ces modèles permettent d'assouplir la vision désormais usuelle de la propriété de cointégration. En effet, les modèles linéaires à correction d'erreur¹ supposent qu'à chaque date, les écarts à l'équilibre de long terme — *i.e.* la relation de cointégration — sont en partie corrigés. Or, il est bien souvent raisonnable de supposer que l'ajustement vers la valeur d'équilibre n'a pas lieu à chaque date : pour que l'ajustement ait lieu, encore faut-il qu'il soit profitable pour les agents concernés, et donc que le gain attendu de l'ajustement de la variable de contrôle soit supérieur ou égal aux éventuels coûts associés à cet ajustement. Ainsi, lorsqu'il s'agit de justifier microéconomiquement les rigidités de prix, Mankiw [1985] met en avant l'existence de coûts de catalogue. De même, les coûts afférents aux transactions financières peuvent créer une zone de non-arbitrage autour des prix ou rendements d'équilibre du marché : comme le précise Anderson [1997] par exemple, tant que l'écart de rendement entre deux titres n'excède pas les coûts de transaction, l'arbitrage (et donc l'ajustement vers la valeur d'équilibre) n'est pas rentable². Un dernier exemple déjà très largement exploré empiriquement³ concerne les arbitrages internationaux sur le marché des biens. Si un arbitrage instantané et sans coût était possible, le prix d'un même bien échangé dans deux pays différents devrait être identique une fois converti dans une même monnaie : la loi du prix unique devrait être vérifiée à chaque date. En transposant ce raisonnement à un panier de biens, on obtiendrait la relation de parité des pouvoirs d'achat. Or, les échanges de biens entre pays sont coûteux en termes de frais de transport ou de droits de douane par exemple. Ici encore, il est possible de montrer (Sercu, Uppal et Van Hulle [1995], Bec, Ben Salem et Carrasco [2004]) que si l'écart de prix entre pays n'est pas suffisamment grand pour couvrir l'ensemble de ces coûts, l'échange n'aura pas lieu et la différence de prix persistera. Les

1. Voir Engle et Granger [1987] ou Johansen [1996] pour une démonstration du théorème de la représentation des processus cointégrés.

2. Ce type de non-linéarité dans le spread de taux est également exploré par Enders et Granger [1998], Enders et Siklos [2001], Bec et Rahbek [2002] ou Bec, Guay et Guerre [2003].

3. Voir entre autres Obstfeld et Taylor [1997], Kapetanios et Shin [2002].

coûts peuvent donc être à l'origine d'écart persistants à une relation d'équilibre de long terme. Les modèles SETAR à trois régimes permettent de rendre compte explicitement de ce type d'ajustement discontinu à l'équilibre : les périodes de non arbitrage, ou non ajustement, sont capturées par un régime central où le comportement du processus peut être de type marche aléatoire ou même explosif. Les régimes extérieurs (inférieur et supérieur) définissent au contraire des zones qui ramènent le processus vers l'équilibre de long terme : le processus s'y comporte comme un processus auto-régressif stationnaire.

Bien que la classe des modèles auto-régressifs à seuil soit apparue il y a plus de vingt ans⁴, ses propriétés statistiques sont loin d'être entièrement explorées. Sous l'hypothèse de stationnarité du processus, la distribution des estimateurs des moindres carrés non-linéaires ou du maximum de vraisemblance de ce type de modèles est connue, ce qui permet de réaliser à peu près l'ensemble des tests d'hypothèses utiles au modélisateur⁵. Les problèmes surgissent lorsqu'il est possible que les séries étudiées ne soient pas stationnaires, ce qui est fréquent pour les données macro-économiques ou financières : dans ce cas, les résultats de l'estimation d'un modèle SETAR ne sont pas exploitables en termes d'inférence — test de linéarité par exemple — car on en ignore le comportement asymptotique⁶. Ceci conduit naturellement à la question épineuse du test de stationnarité ou de racine unitaire d'un processus susceptible d'être non linéaire. Si l'on se concentre sur une non-linéarité de type SETAR à trois régimes, nous verrons dans la première section de cet article que les conditions garantissant la stationnarité de ce type de processus font encore l'objet de recherches. L'état actuel des connaissances a néanmoins permis l'élaboration de tests de racine unitaire consistants et puissants contre ce type d'alternative non linéaire. Les chances sont ainsi plus grandes de détecter des processus cointégrés avec ajustement discontinu.

Il est alors naturel de s'interroger sur la contribution à l'ajustement vers l'équilibre de long terme des différentes variables impliquées dans la relation de cointégration. Dans un cadre linéaire, ceci a été rendu possible par le théorème de la représentation de Granger qui démontre qu'un processus multivarié cointégré admet une représentation vectorielle à correction d'erreur (VECM). Existe-t-il une représentation analogue appropriée à un processus cointégré suivant une dynamique de type SETAR ? Le modèle SETAR étant linéaire par morceaux, la tentation est grande d'extrapoler les résultats de la théorie économétrique des modèles vectoriels auto-régressifs linéaires⁷. Nous verrons toutefois que la discontinuité de ce type de processus générateur de données invalide bien souvent ce raisonnement simple. La section 2 présentera les premiers résultats théoriques formulés récemment

4. Voir l'article fondateur de Tong et Lim [1980].

5. Voir par exemple Granger et Terasvirta [1993].

6. Bien que s'inscrivant dans un modèle un peu différent du SETAR, l'analyse de Caner et Hansen [2001] illustre bien l'interaction non-stationnarité/non-linéarité dans la conception et la distribution asymptotique des tests correspondants.

7. Voir par exemple l'application à la structure par terme d'un modèle vectoriel à correction d'erreur avec ajustement discontinu de Clements et Galvao [2003].

en matière de représentation vectorielle des processus cointégrés avec ajustement discontinu.

Enfin, la section 3 propose une application de ces développements récents de l'économétrie non-linéaire à la structure par terme des taux d'intérêt américains. L'étude empirique, menée sur la base de données hebdomadaires couvrant la période 1960-2003, illustre la pertinence des modèles vectoriels à correction d'erreur à ajustement discontinu.

1. La stationnarité d'un processus SETAR

Dans cette section, nous étudions le modèle SETAR(p) – où p désigne le nombre de retards auto-régressifs – qui est approprié pour l'analyse d'un processus de dimension 1. Il peut s'agir d'une série temporelle observée, X_t , ou d'une série des écarts à une relation d'équilibre de long terme supposée connue, par exemple $X_t = X_{1,t} - \beta_0 X_{2,t}$ si la relation de cointégration fait intervenir deux variables, avec le vecteur cointégrant $(1, -\beta_0)'$. En effet, dans ce dernier cas, les développements actuels de la théorie économétrique relative à l'inférence supposent connue la relation de cointégration et ne s'appliquent donc pas au cas où celle-ci est estimée.

1.1. Conditions de stationnarité

Alors que l'analyse des processus linéaires AR(p), avec $p > 1$, constitue une extension relativement simple de celle des processus AR(1), l'introduction de plus d'un retard dans le modèle SETAR complique considérablement l'analyse et les résultats. Nous nous proposons donc de commencer cette section avec l'étude d'un processus SETAR(1) univarié, X_t , défini par :

$$\Delta X_t = \begin{cases} \pi X_{t-1} + \varepsilon_t & \text{si } |X_{t-1}| > \lambda \\ \varepsilon_t & \text{si } |X_{t-1}| \leq \lambda \end{cases} \quad [1]$$

où π est un paramètre réel constant, λ est le paramètre de seuil réel positif, et ε_t est une séquence de variables aléatoires indépendantes et identiquement distribuées. Ce modèle suppose que les seuils définissant les trois zones sont symétriques : le régime central est défini par $\lambda_{\text{inf}} \leq X_{t-1} \leq \lambda_{\text{sup}}$, où $\lambda_{\text{sup}} = -\lambda_{\text{inf}} = \lambda$. Cette hypothèse, qui a l'avantage de simplifier considérablement la présentation des modèles et des résultats sans perte de géné-

ralité, sera maintenue tout au long de l'article⁸. En contrepartie, nous supposerons désormais que les séries sont centrées autour de zéro. Une deuxième hypothèse simplificatrice concerne la variable gouvernant les changements de régime. De façon générale, il suffit que cette variable soit pré-déterminée, donc retardée de d périodes, avec $d \geq 1$. Le problème est alors d'estimer ce paramètre de délai, d . Comme le remarquent Bec, Ben Salem et Carrasco [2004] sur la base d'exercices de simulations, l'estimateur de ce paramètre est particulièrement peu fiable pour des séries exhibant un certain degré de persistance comme c'est le cas pour la plupart des agrégats qui nous intéressent⁹. Pour cette raison, nous privilégierons ces considérations d'ordre pratique en maintenant $d = 1$ par la suite, au prix d'une perte de généralité.

Le modèle (1) peut également s'écrire :

$$\Delta X_t = s_t \pi X_{t-1} + \varepsilon_t \quad [2]$$

où s_t est la fonction de transition entre les régimes. Ici, il s'agit de la fonction indicatrice qui vaut 1 si $|X_{t-1}| > \lambda$ et 0 sinon. Le processus décrit en (1) ou en (2) a ceci de particulier qu'il se comporte comme une marche aléatoire dans la zone centrale correspondant à de « petites » réalisations de X_{t-1} en valeur absolue, *i.e.* $|X_{t-1}| \leq \lambda$. Comme l'ont montré Chan, Petrucci, Tong et Woolford [1985], le processus X_t défini en (2) est géométriquement ergodique si et seulement si $\pi \in]-2, 0[$. On peut alors trouver une distribution initiale X_0 telle que X_t soit stationnaire. Dans le modèle plus général suivant :

$$\Delta X_t = s_t \pi X_{t-1} + (1 - s_t) \Pi X_{t-1} + \varepsilon_t \quad [3]$$

où Π est un paramètre réel qui n'est plus nécessairement égal à zéro, la condition $\pi \in]-2, 0[$ constitue encore une condition nécessaire et suffisante à l'ergodicité de X_t . Peu importent donc les propriétés du régime central, qui peut être caractérisé par une dynamique stationnaire, ou marche aléatoire ou même explosive : dans ce modèle simple, la stationnarité des régimes extérieurs garantit la stationnarité de l'ensemble du processus¹⁰.

Malheureusement, ce modèle SETAR(1) ne présente qu'un intérêt modéré pour l'étude empirique des problématiques évoquées dans l'introduction. En effet, l'expérience montre que lorsque l'on travaille avec des séries macro-économiques ou financières, il est souvent nécessaire d'introduire un plus grand nombre de retards dans la dynamique afin d'éliminer l'auto-

8. Tous les résultats présentés par la suite s'étendent aisément au cas de seuils asymétriques.

9. Dans ce cas en effet, $X_{t-1}, X_{t-2}, X_{t-3} \dots$ évoluant de façon très similaire, la fonction de vraisemblance est très plate par rapport à d ce qui affecte considérablement la précision de l'estimation.

10. Des conditions suffisantes de stationnarité sont obtenues par Cline et Pu [1999] dans le cas SETAR(1) avec un délai d supérieur ou égal à 1.

corrélation des résidus estimés. Le SETAR(p) suivant est donc un modèle plus utilisé :

$$\Delta X_t = s_t [\pi X_{t-1} + \Gamma(L) \Delta X_{t-1}] + (1 - s_t) [\Pi X_{t-1} + G(L) \Delta X_{t-1}] + \varepsilon_t \quad [4]$$

mais la condition obtenue par Chan *et al.* [1985] ne garantit plus la stationnarité de X_t . Elle n'est même pas suffisante. Tjøstheim [1990], qui retrouve cette condition dans le cas $p = 1$, ne parvient à l'étendre au cas $p \geq 1$ qu'en faisant dépendre les changements de régime d'une norme des p retards auto-régressifs. Plus précisément, dans le cas où $s_t = 1_{\{\|x_{t-1}, \dots, x_{t-p}\| > \lambda\}}$, Tjøstheim [1990] montre qu'une condition suffisante est que la plus petite valeur propre en valeur absolue du polynôme caractéristique associé au régime extérieur soit inférieure à un. A ce jour, aucune condition nécessaire et suffisante de stationnarité n'est disponible pour le modèle (4) lorsque s_t ne dépend que de X_{t-1} . Chan et Tong [1985] ont montré qu'une condition suffisante à l'ergodicité d'un tel processus est que la dynamique de chaque régime soit stationnaire. Bec, Ben Salem et Carrasco [2004] assouplissent ce résultat en montrant que sous des hypothèses raisonnables, un processus tel que celui décrit en (4) avec une racine unitaire dans le régime central est encore stationnaire, à condition que la plus petite valeur propre en valeur absolue du polynôme caractéristique associé au régime extérieur soit inférieure à un¹¹. Même s'il ne s'agit « que » d'une condition suffisante, ce résultat est nécessaire pour l'élaboration, et l'étude du comportement asymptotique, de tests de racine unitaire adaptés à ce type de processus.

1.2. Tests de racine unitaire

Selon Balke et Fomby [1997], la non linéarité du processus (4) ne devrait pas empêcher les tests standards de racine unitaire de détecter la stationnarité de X_t le cas échéant, ce que semblent confirmer les simulations présentées dans cette étude. De nombreux travaux, tels que ceux de Pippenger et Goering [1993], Pippenger et Goering [2000], ou encore Taylor [2001], viennent toutefois nuancer ce résultat en remettant en cause la puissance du test ADF face à une alternative stationnaire de type SETAR. Il ressort principalement de ces études que le test ADF ne rejette pas assez souvent l'hypothèse nulle de racine unitaire *i)* lorsque le seuil est grand et *ii)* lorsque beaucoup d'observations sont situées dans une zone centrale de type marche aléatoire. Les exercices de simulation présentés dans Bec, Ben Salem et Carrasco [2004] et Bec, Guay et Guerre [2003] confirment le manque de puissance du test ADF face à ce type d'alternatives qui semblent pourtant caractériser la dynamique de variables telles que les taux de change réels des pays du G7 ou encore les spreads taux long - taux court de pays tels que

11. Dans le cas où la variable de transition serait X_{t-d} avec $d \geq p$, ces auteurs montrent qu'une racine explosive dans le régime central ne remet pas en cause la stationnarité de X_t , pourvu que les régimes extérieurs soient stationnaires.

la France, l'Allemagne ou les États-Unis¹². Il est donc souhaitable de recourir à des tests de racine unitaire spécifiquement conçus pour une hypothèse alternative de type SETAR. Le test proposé par Bec, Ben Salem et Carrasco [2004] consiste à tester l'hypothèse nulle

$$H_0 : \rho_1 = \rho_2 = \rho_3 = 0$$

dans le SETAR(p) suivant :

$$\Delta X_t = \begin{cases} a_{11} \Delta X_{t-1} + \dots + a_{1p-1} \Delta X_{t-p+1} + \mu_1 + \rho_1 X_{t-1} + \varepsilon_t & \text{if } X_{t-1} \leq -\lambda \\ a_{21} \Delta X_{t-1} + \dots + a_{2p-1} \Delta X_{t-p+1} + \mu_2 + \rho_2 X_{t-1} + \varepsilon_t & \text{if } |X_{t-1}| < \lambda \\ a_{31} \Delta X_{t-1} + \dots + a_{3p-1} \Delta X_{t-p+1} + \mu_3 + \rho_3 X_{t-1} + \varepsilon_t & \text{if } X_{t-1} \geq \lambda \end{cases}$$

qui est plus général que le modèle (4) car il autorise une dynamique auto-régressive différente dans les deux régimes extérieurs. La statistique de test est¹³

$$\sup LR \equiv \sup_{\lambda \in [\underline{\lambda}_T, \bar{\lambda}_T]} LR_T(\lambda)$$

où les bornes $\underline{\lambda}_T$ et $\bar{\lambda}_T$ sont choisies de façon à s'assurer qu'il reste au moins 10 % des observations dans le régime central et dans les régimes extérieurs. $LR_T(\lambda)$ est la statistique du rapport de vraisemblance correspondant au seuil λ définie par :

$$LR_T(\lambda) \equiv T \ln \left(\frac{\tilde{\sigma}^2(\lambda)}{\hat{\sigma}^2(\lambda)} \right)$$

où $\tilde{\sigma}^2(\lambda)$ et $\hat{\sigma}^2(\lambda)$ désignent la variance des résidus estimés des modèles contraint et non contraint respectivement. Les simulations réalisées pour des valeurs des paramètres du PGD comparables à celles estimées pour le taux de change réel des principaux pays industrialisés révèlent que la puissance de ce test Sup LR, de l'ordre de 90 %, est trois fois plus grande que celle du test ADF.

12. Le résultat des simulations présentées dans Balke et Fomby [1997] s'explique par la valeur du paramètre auto-régressif retenue par ces auteurs dans le régime extérieur, et qui correspond à $\pi = -0,6$ dans le modèle (1) : le PGD sous l'alternative est donc très éloigné d'un processus racine unitaire.

13. On peut de la même façon retenir une statistique SupWald ou Sup LM.

Bec, Guay et Guerre [2003] proposent une version modifiée de ce test visant à le rendre adaptatif : le choix de l'ensemble de seuils possibles $\Lambda_T = [\underline{\lambda}_T, \bar{\lambda}_T]$ s'adapte selon que l'on se place sous l'hypothèse nulle ou sous l'hypothèse alternative. En effet, pour accroître la puissance du test, il est souhaitable que les valeurs critiques ne soient pas trop élevées, et donc que l'ensemble des seuils balayés pour calculer la statistique ne soit pas trop grand sous l'hypothèse nulle. Au contraire, pour que la valeur de la statistique soit élevée sous l'hypothèse alternative, il faut se donner toutes les chances que l'ensemble de seuils contienne la vraie valeur du seuil : il faut donc balayer sur l'ensemble le plus large possible. L'idée ici est de faire dépendre les bornes $\underline{\lambda}_T$ et $\bar{\lambda}_T$ de la statistique ADF¹⁴, celle-ci étant d'autant plus élevée que le processus est loin de la racine unitaire. Ainsi, en choisissant $\underline{\lambda}_T$ (resp. $\bar{\lambda}_T$) comme une fonction décroissante (resp. croissante) de la statistique ADF, $\hat{\tau}_T$, on obtient un ensemble Λ_T plus grand sous l'alternative que sous la nulle. En notant $\hat{\sigma}_{0T}$ l'écart-type estimé des résidus de la régression du test ADF, $|X|_{(3)}$ la troisième observation de la série des valeurs absolues de $\{X_{t-1}\}_{t=1}^T$ classées par ordre croissant, et l un paramètre fixé à 6 pour optimiser la puissance du test, l'intervalle adaptatif est défini par :

$$\underline{\lambda}_T = |X|_{(3)} + \frac{\hat{\sigma}_{0T}}{\max(1, |\hat{\tau}_T|)} \quad \text{et} \quad \bar{\lambda}_T = \underline{\lambda}_T + l\hat{\sigma}_{0T} \max(1, |\hat{\tau}_T|)$$

Les simulations révèlent que cette correction améliore sensiblement la puissance du test de racine unitaire. En outre, cette statistique de test est consistante, *i.e.* elle diverge, contre toute alternative stationnaire ergodique, linéaire ou non¹⁵. C'est donc ce test que nous privilégierons par la suite.

Ces développements récents nous fournissent ainsi des tests de racine unitaire plus puissants contre les alternatives stationnaires non linéaires que le test ADF.

2. Le modèle VECM à seuil (TVECM)

Lorsqu'ils sont appliqués à une série d'écart observés à une relation d'équilibre, les tests présentés plus haut sont en mesure de détecter un processus cointégré avec ajustement discontinu tel que ceux évoqués dans l'introduction. Du point de vue de la théorie économique comme du point de vue empirique — pour améliorer la prévision par exemple — il est alors au moins aussi intéressant de savoir comment les séries impliquées dans la

14. Ou de n'importe quelle autre statistique de test de racine unitaire consistante contre l'alternative stationnaire.

15. Le test développé par Bec, Ben Salem et Carrasco [2004] est consistant contre l'alternative SETAR stationnaire.

relation de cointégration réagissent aux déviations à cette relation. Dans l'analyse linéaire, le modèle VECM est parfaitement adapté à cette question. Qu'en est-il lorsque l'ajustement est discontinu ?

2.1. La représentation de processus cointégrés avec ajustement discontinu

Quelques articles récents se sont penchés sur l'ergodicité géométrique de modèles vectoriels autorisant un ajustement non linéaire à une combinaison linéaire stationnaire des variables considérées. Les résultats relatifs à l'ergodicité présentés dans Escribano et Mira [2002] et Corradi, Swanson et White [2000] s'appliquent lorsque la fonction gouvernant les changements de régime, s_t , est convenablement lisse, continue et mesurable. Ceux de Saikkonen [2002] s'appliquent dans le cas discontinu, mais pas dans le cas d'une fonction de transition non observable. Ces deux derniers cas sont traités dans Bec et Rahbek [2002].

Considérons pour commencer un processus X_t de dimension 2, satisfaisant les conditions du théorème 4 de Johansen [1996]. Ceci implique que X_t est intégré d'ordre 1, alors que ΔX_t et la relation de cointégration βX_t sont stationnaires. Supposons maintenant que l'ajustement vers βX_t ne soit pas linéaire, au sens où il ne serait pas opérant à chaque date, comme par exemple dans l'analogie vectoriel du modèle (1) :

$$\Delta X_t = \begin{cases} \alpha \beta X_{t-1} + \varepsilon_t & \text{si } |\beta X_{t-1}| > \lambda \\ \varepsilon_t & \text{si } |\beta X_{t-1}| \leq \lambda \end{cases} \quad [5]$$

Selon le modèle (5), seuls les écarts à la relation de cointégration supérieurs à λ en valeur absolue sont corrigés via le vecteur de paramètres α . Ceci correspond précisément au comportement d'arbitrage en présence de coûts de transaction évoqué dans l'introduction. Les conditions garantissant la stationnarité du processus défini en (5) doivent toutefois être spécifiées et vérifiées. Le théorème 2 de Bec et Rahbek [2002] explicite ces conditions pour le modèle TVECM — Threshold Vector Error Correction Model — plus général suivant :

$$\Delta X_t = s_t \left[\alpha \beta X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} \right] + (1 - s_t) \left[\alpha \beta X_{t-1} + \sum_{i=1}^{p-1} G_i \Delta X_{t-i} \right] + \varepsilon_t \quad [6]$$

où X_t est de dimension n , α , α et β sont de dimension $n \times r$, $(\Gamma_i)_{i=1, \dots, p-1}$ et $(G_i)_{i=1, \dots, p-1}$ sont des matrices $n \times n$. La séquence ε_t est supposée i.i.d. $(0, \Omega)$ avec $\Omega > 0$, et les $(X_{-i})_{i=0, \dots, k-1}$ sont fixés. La fonction de transition est définie comme

$$s_t = p(x_t) = 1 (\|x_t\| \geq \lambda) \quad [7]$$

avec

$$x_t = (X'_{t-1}\beta, \Delta X'_{t-1}, \dots, \Delta X'_{t-p+1})' \quad [8]$$

Ce théorème établit notamment que si X_t est défini par le système (6), (7) et (8), et si le polynôme caractéristique des régimes extérieurs

$$A(z) = I_n(1-z) - \alpha\beta z - \sum_{i=1}^{p-1} \Gamma_i(1-z)z^i$$

a $n-r$ racines en $z=1$ et les r autres telles que $|z| > 1$, alors X_t est non-stationnaire non-linéaire, βX_t et ΔX_t sont stationnaires, et tous leurs moments existent et sont finis. Il est intéressant de remarquer que la condition sur les racines du polynôme caractéristique donnée ci-dessus est analogue à celle de Johansen [1996], mais elle ne s'applique ici qu'à la dynamique du régime extérieur, correspondant aux grandes déviations à l'équilibre de long terme. Une dynamique non-stationnaire ou même explosive du régime central ne remet pas en cause ce résultat¹⁶. Toutefois, ce dernier est obtenu pour une fonction de transition s_t qui dépend de x tel que défini en (8), donc non seulement des déséquilibres mesurés par βX_{t-1} mais aussi des différences premières ΔX_{t-i} . Or, le cas où s_t ne dépend que des écarts à l'équilibre, *i.e.*

$$s_t = 1 (\|\beta X_{t-1}\| \geq \lambda) \quad [9]$$

est sans doute plus approprié au regard des comportements d'arbitrage dont nous souhaitons rendre compte¹⁷. Le théorème 3 de Bec et Rahbek [2002] établit que dans le cas où s_t est défini par (9), les résultats du théorème 2 énoncés plus haut restent valides si les paramètres de court terme des régimes extérieurs et central sont identiques, donc pour le modèle :

$$\Delta X_t = [s_t \alpha + (1-s_t)a] \beta X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \varepsilon_t \quad [10]$$

2.2. Estimation et tests

Bec et Rahbek [2002] montrent que si β et λ sont connus et fixés, alors les estimateurs du maximum de vraisemblance des paramètres du modèle (6), et donc *a fortiori* du modèle (10), sont convergents et distribués asymptotiquement selon une loi Normale. Cependant, alors que la théorie économique nous renseigne parfois sur la valeur des paramètres des vecteurs

16. En effet, comme la probabilité que $s_t = 1$ tend vers 1 lorsque $\|x\| \rightarrow \infty$, le processus ne peut être absorbé dans le régime central.

17. Voir par exemple l'application de Clements et Galvao [2003].

cointégrants, ce n'est pas le cas pour la valeur du seuil : l'évaluation des coûts de transaction est de nature empirique. Il faut donc estimer le seuil λ . Hansen et Seo [2002] proposent de l'estimer en balayant sur un ensemble de valeurs possibles, $\Lambda = [\lambda_L, \lambda_U]$, dont les bornes sont choisies de façon à laisser au moins 5 % des observations dans chaque régime. Ainsi, pour chaque $\lambda \in \Lambda$, on obtient les estimateurs des moindres carrés non-linéaires de $\hat{\alpha}(\lambda)$, $\hat{\beta}(\lambda)$, $\hat{\Gamma}_i(\lambda)$, $\hat{\Omega}(\lambda)$ ainsi que la vraisemblance correspondante, $L(\lambda)$, dans le modèle (9)-(10) que nous retiendrons par la suite¹⁸. L'estimateur du seuil, $\hat{\lambda}$, est alors la valeur qui maximise la fonction de vraisemblance.

Les résultats de Bec et Rahbek [2002] relatifs à la convergence et à la distribution asymptotique des estimateurs du TVECM n'ont pas encore été étendus au cas où le seuil est estimé. Néanmoins, Hansen et Seo [2002] font la conjecture que cette estimation du maximum de vraisemblance devrait produire des estimateurs convergents de distribution asymptotiquement normale, comme si le seuil était connu.

Il est alors possible de tester l'hypothèse nulle de linéarité de l'ajustement à l'équilibre contre l'alternative d'un ajustement à seuil, en remarquant que le VECM linéaire est un cas particulier du TVECM que l'on obtient en posant l'égalité des coefficients α_i et a_i entre régimes. Dans le modèle (10), l'hypothèse nulle de linéarité s'écrit donc :

$$H_0 : \alpha_i = a_i, \quad i = 1, \dots, n \quad [11]$$

Bec, Ben Salem et Rahbek [2003] développent la statistique de test correspondante, en reformulant le modèle (10) comme suit :

$$Y_t = BZ_t + \varepsilon_t$$

où $Y_t = \Delta X_t$ est un vecteur $n \times 1$, $Z_t = (s_t \beta' X_{t-1}, (1-s_t) \beta' X_{t-1}, \Delta X_{t-1}, \dots, \Delta X_{t-p+1})'$ est un vecteur $l \times 1$ et B' est la matrice $n \times l$ de paramètres telle que l'hypothèse nulle se réécrit

$$H_0 : B = H\Psi \quad [12]$$

avec H et Ψ de dimension $l \times s$ et $s \times n$ respectivement, et $s < l$. La statistique du multiplicateur de Lagrange de l'hypothèse nulle (12) est donnée par :

$$LM(\lambda) = T \operatorname{tr} \{ S_{ZZ}^{-1} H_{\perp} (H'_{\perp} S_{ZZ}^{-1} H_{\perp})^{-1} H'_{\perp} S_{ZZ}^{-1} S_{ZY} \hat{\Omega}^{-1} S_{YZ} \} \quad [13]$$

où T est le nombre d'observations, $S_{ZZ} = \frac{1}{T} \sum_1^T Z_t Z_t'$, $S_{YZ} = \frac{1}{T} \sum_1^T Y_t Z_t'$ et H_{\perp} est le complément orthogonal de H , de dimension $l \times (l-s)$. Cette statistique LM dépend de λ via les deux premières variables de Z_t . Si β et λ étaient

18. La façon de procéder serait identique si le modèle étudié était constitué des équations (6)-(8).

connus, elle serait distribuée selon un $\chi_n^2(1-s)$. Si le seuil est inconnu, λ est un paramètre de nuisance car on ne peut l'estimer sous l'hypothèse nulle de linéarité : on ne peut donc pas calculer cette statistique. La solution proposée par Hansen [1996] consiste à recourir à la statistique Sup LM suivante :

$$\text{Sup LM} = \sup_{\lambda \in \Lambda} \text{LM}(\lambda) \quad [14]$$

Comme la distribution asymptotique non-standard de cette statistique dépend de fonctions de moments inconnues, elle ne peut être tabulée. Les p-values peuvent cependant être obtenues par une méthode de bootstrap des résidus comme celle développée par Hansen [1996]¹⁹. Dans l'application qui suit, elles seront calculées à partir de 1000 simulations.

3. Une application : la structure par terme des taux d'intérêt

Nous nous proposons maintenant d'appliquer les développements précédents au test de l'hypothèse de la théorie des anticipations de la structure par terme des taux d'intérêt.

Lorsque les ajustements de portefeuille sont instantanés et en l'absence de coûts de transaction, cette théorie implique qu'à l'équilibre le spread de taux, $S_t(1, k)$, est stationnaire²⁰. Elle implique donc l'existence d'une relation de cointégration entre les taux d'intérêt de long terme et court terme, le vecteur cointégrant étant donné par $\beta = (1, -1)'$. De très nombreux travaux empiriques ont été consacrés au test de cette première implication de la théorie des anticipations de la structure par terme. Les résultats sont assez mitigés, et dépendent clairement de l'échantillon de données retenu, *i.e.* de la période couverte, du pays considéré, mais aussi des termes choisis pour les taux d'intérêt candidats à la cointégration. Par exemple, Campbell et Shiller [1987], Stock et Watson [1988], Anderson [1997] ou encore Tzavalis et Wickens [1998] concluent à la stationnarité du spread américain alors que les études de Evans et Lewis [1994], Enders et Siklos [2001] ou Bec, Guay et Guerre [2003] ne rejettent pas l'hypothèse de racine unitaire pour cette variable.

Toutefois, comme le remarque Anderson [1997], il est tout à fait probable que l'ajustement à cette relation de long terme ne soit pas linéaire. En effet, si l'on tient compte de l'existence de coûts de transaction, un investisseur ne modifiera la composition de son portefeuille d'actifs que si le gain qu'il espère retirer de cette opération fait plus que compenser le coût qu'elle engendre. Considérons des coûts de transaction homogènes, qui viennent

19. Voir Hansen et Seo [2002] ou Bec, Ben Salem et Rahbek [2003] pour plus de détails.

20. Voir Anderson [1987] par exemple.

réduire le rendement du titre acheté par l'investisseur d'un montant λ donné et constant. Dans ce cas, l'investisseur convertira son portefeuille de titres à une période en titres à k périodes si et seulement si $\lambda < S_t(1, k) - \hat{S}_t(1, k)$, où le membre de droite de l'inégalité est le spread observé en écart à sa valeur d'équilibre. De même, il échangera ses titres longs contre des titres courts si et seulement si $S_t(1, k) - \hat{S}_t(1, k) < -\lambda$. Par conséquent, l'introduction de coûts de transaction homogènes implique l'existence d'une zone de non-arbitrage définie par : $-\lambda < S_t(1, k) - \hat{S}_t(1, k) < \lambda$. L'ajustement à l'équilibre de long terme n'est pas continu. Il n'intervient que lorsque la déviation à l'équilibre de long terme est grande²¹. Le modèle TVECM (9)-(10) présenté dans la section précédente permet de rendre compte d'une telle dynamique discontinue d'ajustement des taux d'intérêt, avec : $X_t = (R_t^L, R_t^S)$ où R_t^L et R_t^S sont les taux d'intérêt long et court respectivement, $\beta = (1, -1)$ et $s_t = 1_{(|S_t(1, k) - \hat{S}_t(1, k)| \geq \lambda)}$.

Ainsi que Campbell et Shiller [1987] l'ont souligné, une autre implication de la théorie des anticipations de la structure par terme est que comme le spread est une combinaison linéaire des variations futures anticipées du taux d'intérêt de court terme, il devrait contribuer à la prévision du taux court, *i.e.* le spread devrait causer au sens de Granger le taux d'intérêt de court terme. Dans le modèle TVECM (9)-(10), on s'attend à ce que cette influence ne s'exerce que dans les régimes extérieurs, donc pour de grandes valeurs du spread en valeur absolue.

3.1. Les données

Notre application porte sur les rendements nominaux à maturité des bons du Trésor américain à trois et six mois depuis 1960. S'agissant de données hebdomadaires²², nous disposons donc de 2 047 observations représentées dans le graphique 1. Il s'agit de rendements de bons nouvellement émis, dont l'adjudication est effectuée une fois par semaine par le Département du Trésor américain. Ces bons du Trésor à trois et six mois sont par conséquent identiques en tous points, excepté le temps restant jusqu'à échéance. L'avantage de ces données de taux est qu'elles reflètent le prix auquel se font effectivement les transactions sur le marché primaire des bons du Trésor américain : il s'agit du prix moyen des offres des *primary dealers* calculé le jour de l'adjudication. Les principaux coûts de transaction sur ce marché particulier incluent les coûts de recherche des *dealers* ainsi que les coûts associés à la revente de ces titres sur le marché secondaire.

21. En testant la racine unitaire contre une alternative stationnaire non linéaire, Enders et Siklos [2001] et Bec, Guay et Guerre [2003] rejettent l'hypothèse nulle non stationnaire. De façon plus générale, les analyses de Anderson [1997], Enders et Granger [1998] ou encore Bohl et Siklos [2001] concluent en faveur d'un tel comportement non linéaire du spread.

22. Source : Federal Reserve Bank of St. Louis.

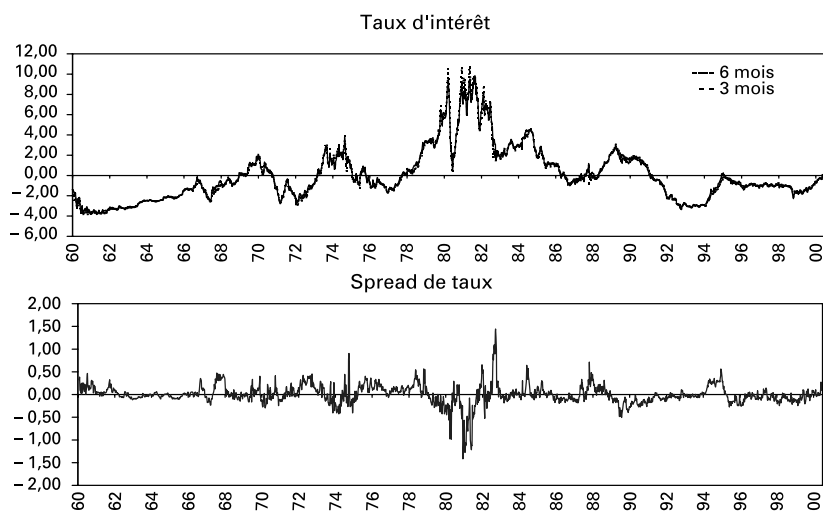


Figure 1. Les données.

Bien qu'indispensable à l'analyse d'un ajustement à l'équilibre de long terme, la longueur de l'échantillon étudié ici est toutefois susceptible de capter des ruptures dans la conduite de la politique monétaire américaine. En effet, on peut distinguer au moins trois sous-périodes depuis 1960. La première s'arrête en septembre 1979 et correspond à une période où la Federal Reserve Bank a mené une politique visant à maintenir le taux des fonds fédéraux dans une fourchette cible spécifiée par le Federal Open Market Committee. Ensuite, entre octobre 1979 et septembre 1982, la Fed a mis en place de « nouvelles procédures opératoires »²³ consistant à contrôler le niveau des réserves non empruntées de façon à atteindre l'objectif de masse monétaire défini par le FOMC. L'abandon de l'objectif de taux au cours de ces trois années se traduit par une plus grande volatilité des taux des bons du Trésor étudiés ici, comme le montre le graphique 1. Depuis octobre 1982, la politique monétaire cherche à contrôler le niveau des réserves empruntées. Ainsi que le remarque Gilbert [1985], cette procédure revient à contrôler indirectement le taux des fonds fédéraux et les autres taux de court terme : comme dans la politique menée jusqu'en septembre 1979, l'offre de réserves non empruntées répond aux variations de la demande de réserves de telle sorte que ces variations affectent peu le federal funds rate. Afin d'évaluer l'influence de ces changements structurels sur nos résultats, outre l'étude sur l'ensemble de l'échantillon (T) incluant toutes les observations hebdomadaires de la première semaine de janvier 1960 à la dernière semaine de juin 2003, nous mènerons également l'analyse en distinguant les trois sous-périodes suivantes :

(T1) 1^{re} semaine de janvier 1960 - 4^e semaine de septembre 1979,

23. Il s'agit des « New operating procedures » mises en place sous la direction de Paul Volker.

(T2) 1^{er} semaine d'octobre 1979 - 4^e semaine de septembre 1982,

(T3) 1^{er} semaine d'octobre 1982 - 4^e semaine de juin 2003.

3.2. Tests préliminaires

Dans la perspective de l'estimation du modèle non linéaire, nous travaillons désormais sur les variables centrées. Ceci présente l'avantage de considérer une bande d'inaction centrée autour de zéro ce qui permet d'estimer un seul seuil, le deuxième étant simplement son opposé. Les résultats des tests ADF appliqués à R_t^S et R_t^L ne permettent pas de rejeter l'hypothèse nulle de racine unitaire pour le processus des taux d'intérêt à 3 et 6 mois, sur l'ensemble de la période comme sur les trois sous-périodes. Au contraire, comme le montrent les résultats reportés dans le tableau 1, la racine unitaire est très clairement rejetée pour le spread de taux selon le test ADF aussi bien que selon le test Sup Wald (Λ^T) de Bec, Guay et Guerre [2003], quelle soit la période considérée. Ces conclusions sont en grande partie confirmées par les tests de cointégration de Johansen [1996] effectués à partir du système bivarié (R_t^L, R_t^S). Les résultats des tests de la Trace et du $\hat{\lambda}_{\max}$ (voir le tableau 3 en annexe) permettent de conclure à l'existence d'une relation de cointégration²⁴, la restriction $\beta = (1, -1)$ n'étant pas rejetée au seuil de 5 %, pour les échantillons (T), (T1) et (T3), confortant ainsi les implications de la théorie des anticipations. Pour la sous-période (T2), qui couvre les trois années très perturbées, il n'existe pas de relation de cointégration selon ce test.

Tableau 1. Tests de racine unitaire pour S_t

Période étudiée	nombre d'obs.	nombre de retards	stat ADF	p-val.	stat Sup Wald (Λ^T)	p-val.
(T)	2040	9	- 7,62	1 %	65,71	1 %
(T1)	1030	2	- 7,18	1 %	62,34	1 %
(T2)	156	2	- 2,67	1 %	28,72	1 %
(T3)	861	2	- 6,71	1 %	56,95	1 %

24. Le nombre de retards du modèle est choisi de façon à éliminer l'auto-corrélation des résidus dans le système en différences premières. Les conclusions sont peu sensibles au choix du nombre de retards.

3.3 Estimation du TVECM

Ainsi que nous l'avons souligné plus haut, la théorie des anticipations de la structure par terme des taux d'intérêt implique que le spread de taux doit causer au sens de Granger le taux court. Le test de cette hypothèse à partir du modèle (9)-(10) implique des restrictions assez compliquées sur les paramètres estimés. Il est beaucoup plus simple à réaliser si l'on réécrit le mo-

dèle en $Y_t = MX_t = (S_t, R_t^S)'$, où $M = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, soit encore :

$$\Delta Y_t = \Gamma(L) \Delta Y_{t-1} + [(1 - s_t) \bar{a} + s_t \bar{a}] \tilde{\beta} Y_{t-1} + u_t \quad [15]$$

avec $\Gamma(L) = MG(L)M^{-1}$, $\bar{a} = M\alpha$, $\tilde{a} = Ma$ et $\tilde{\beta} = \beta M^{-1}$. Nous obtenons donc le système d'équations suivant :

$$\begin{cases} \Delta S_t = \Gamma_{11}(L) \Delta S_{t-1} + \Gamma_{12}(L) \Delta R_{t-1}^S + s_t \bar{a}_1 S_{t-1} + \\ \quad (1 - s_t) \bar{a}_1 S_{t-1} + u_{1,t} \\ \Delta R_t^S = \Gamma_{21}(L) \Delta S_{t-1} + \Gamma_{22}(L) \Delta R_{t-1}^S + s_t \bar{a}_2 S_{t-1} + \\ \quad (1 - s_t) \bar{a}_2 S_{t-1} + u_{2,t} \end{cases}$$

Le test de la non-causalité au sens de Granger du spread de taux vers le taux d'intérêt court revient alors à vérifier les contraintes :

$$\begin{cases} \Gamma_{21}(L) = 0 \quad \text{et} \quad \bar{a}_2 = 0, \quad \text{si} \quad s_t = 1 \\ \Gamma_{21}(L) = 0 \quad \text{et} \quad \bar{a}_2 = 0, \quad \text{si} \quad s_t = 0 \end{cases}$$

L'estimation du modèle TVECM ainsi que les statistiques des tests de non-causalité et de linéarité pour l'ensemble de la période sont reportées dans le tableau 2. Comme le montrent les résultats de l'estimation du modèle à seuil, l'ajustement non linéaire à l'équilibre de long terme est validé par les données, puisque l'hypothèse nulle de linéarité est largement rejetée au seuil de 1 %. Le taux court ne réagit aux déviations à l'équilibre de long terme que lorsque celles-ci sont suffisamment importantes : dans les régimes extérieurs, le taux court augmente afin de réduire le spread de taux. En outre, le spread de taux cause le taux court à court terme comme à long terme, conformément à la théorie des anticipations. L'analyse en sous-périodes (voir le tableau (4) en annexe) confirme ces conclusions pour (T1) et (T3).

Tableau 2. Estimation du modèle à seuil pour la période (T)

$\Delta S_t = \Gamma_{11}(L) \Delta S_{t-1} + \Gamma_{12}(L) \Delta R_{t-1}^S - 0,11 s_t S_{t-1} - 0,05 (1 - s_t) S_{t-1}$
$F_{11} = 0,00 \quad F_{12} = 0,00 \quad (-7,91) \quad (-2,54)$
$\Delta R_t^S = \Gamma_{21}(L) \Delta S_{t-1} + \Gamma_{22}(L) \Delta R_{t-1}^S + 0,17 s_t S_{t-1} - 0,00 (1 - s_t) S_{t-1}$
$F_{21} = 0,00 \quad F_{22} = 0,00 \quad (4,70) \quad (-0,01)$
$k = 10; L = -11\ 074,9; \text{seuil} = 0,31; Q_1(12) = 0,22; Q_2(12) = 0,64$
<p>TESTS DE NON CAUSALITÉ :</p> <p>$s_t = 1 : F(10, 1997) = 11,29$ (p - val = 0,00)</p> <p>$s_t = 0 : F(10, 1997) = 6,28$ (p - val = 0,00)</p>
<p>TEST DE LINÉARITÉ :</p> <p>Sup LM = 70,35 (bootstrap p - val < 0,001)</p>

Les statistiques de Student sont données entre parenthèses.
 F_{ij} = p - value de l'hypothèse nulle que tous les coefficients de Γ_{ij} sont égaux à 0.
 $Q_i(h)$ = p - value des tests d'autocorrélation des résidus à l'ordre h dans l'équation i.

Les résultats sont moins concluants pour la période intermédiaire (T2), durant laquelle la Fed a privilégié un objectif en termes de masse monétaire et non de taux. En effet, bien que l'ajustement à l'équilibre de long terme soit encore non linéaire comme on peut le vérifier d'après l'équation du spread, et que ce dernier cause le taux court, le coefficient d'ajustement du taux court à l'équilibre de long terme n'est pas significativement différent de zéro. L'ajustement a donc été réalisé par le taux long. Ces résultats doivent néanmoins être considérés avec précaution, car l'hypothèse de linéarité n'est rejetée qu'au seuil de 34 %, et la période (T2) est à la fois très courte et très volatile.

*
* * *

Les modèles que nous avons traités dans cet article ne sont qu'une infime parcelle d'un vaste champ de recherche. Nous nous proposons pour conclure d'en préciser les contours.

Tout d'abord, la classe de modèles considérée ici est caractérisée par une fonction de transition endogène, ce qui constitue une difficulté majeure. En effet, lorsque la transition est fonction d'une variable exogène stationnaire, cela garantit que chaque régime est visité infiniment souvent et les conditions d'ergodicité du processus modélisé s'en trouvent grandement assouplies (Caner et Hansen [2001] et Tsay [1998] entre autres). Ensuite, la nature discontinue de l'ajustement dans les modèles que nous avons présentés les distingue de la classe voisine des modèles à seuil à transition lisse (modèles STAR). Celle-ci admet l'ajustement discontinu comme cas limite, lorsque l'on fait tendre la vitesse de transition entre régimes vers l'infini. Une littérature analogue à celle que nous avons présentée s'est développée de façon

concomitante pour les modèles STAR (Bec, Ben Salem et Carrasco [2002], Kapetanios, Shin et Snell [2003], Corradi *et al.* [2000], Escribano et Mira [2002], par exemple). Une troisième spécificité des modèles que nous avons présentés concerne le nombre de régimes. Alors que les propriétés de stabilité des modèles à deux régimes sont connues dans un cadre univarié (Chan et Tong [1985]), il n'existe pas à notre connaissance de résultats théoriques équivalents à ceux que nous avons présentés pour les modèles vectoriels²⁵. Une dernière caractéristique des modèles considérés dans cet article est la nature linéaire de la relation de cointégration. Il est tout à fait envisageable que la relation de cointégration soit elle-même non linéaire (Granger [1995]).

Au sein même de cette classe restreinte des modèles *i)* à fonction de transition endogène, *ii)* à ajustement discontinu, *iii)* vers une relation de cointégration linéaire, *iv)* entre trois régimes, l'exploration est loin d'être achevée. En effet, il serait par exemple souhaitable d'étendre les résultats de Bec et Rahbek [2002] au cas où β et λ ne sont pas connus mais estimés par le maximum de vraisemblance. Cette extension est rendue difficile par la non différentiabilité de la fonction de vraisemblance par rapport à ces paramètres.

Références bibliographiques

- ANDERSON H. [1997], Transaction Costs and Non-Linear Adjustment Towards Equilibrium in the U.S. Treasury Bill Market, *Oxford Bulletin of Economics and Statistics*, 59(4), 465-484.
- BALKE N. S. et FOMBY T. B. [1997], Threshold Cointegration, *International Economic Review*, 38, 627-45.
- BEC F., GUAY A. et GUERRE E. [2003], *Adaptive Consistent Unit root Tests Based on Autoregressive Threshold Model*, Version révisée du Working Paper 2002-46, CREST, Malakoff, France.
- et RAHBK A. [2002], *Vector Equilibrium Correction Models with Non-Linear Discontinuous Adjustments*, Preprint 15, Department of Statistics and Operations Research, University of Copenhagen.
- BEN SALEM M. et RAHBK A. [2003], *Nonlinear Adjustment towards The Purchasing Power Parity Relation : A Multivariate Approach*, Mimeo, CREST-ENSAE, Malakoff.
- , — et CARRASCO M. [2002], *The Purchasing Power Parity Puzzle : Evidence from a LSTAR Specification*, Working Paper 2002-47, EUREQua, University of Paris 1.
- , — et — , Test for Unit-root Versus Threshold Specification with an Application to the PPP, *Journal of Business and Economic Statistics*, 2004. A paraître.
- BOHL M. et SIKLOS P. [2001], *The Bundesbank's Inflation Policy and Asymmetric Behavior of the German Term Structure*, manuscript, Dept of economics, European University Viadrina Frankfurt, Germany.

25. L'analyse de Hansen et Seo [2002] n'est pas complète de ce point de vue.

- CAMPBELL J. et SHILLER R. [1987], Cointegration and Tests of Present Value Models, *Journal of Political Economy*, 95(5), 1062-1088.
- CANER M. et HANSEN B. [2001], Threshold Autoregression With a Unit Root, *Econometrica*, 69, 1555-1596.
- CHAN K. S., PETRUCCCELLI J., TONG H. et WOOLFORD S. [1985], A Multiple-Threshold AR(1) Model, *Journal of Applied Probability*, 22, 267-279.
- CHAN K. S. et TONG H. [1985], On the Use of Lyapunov Functions for the Ergodicity of Stochastic Difference Equations, *Advances in Applied Probability*, 17, 666-678.
- CLEMENTS M. P. et GALVAO A. B. [2003], Testing the Expectations Theory of the Term Structure of Interest Rates in Threshold Models, *Macroeconomic Dynamics*, 7, 567-585.
- CLINE D. et PU H. [1999], Stability of nonlinear AR(1) time series with delay, *Stoch. Proc. Appl.*, 82, 307-333.
- CORRADI V., SWANSON N. R. et WHITE H. [2000], Testing for Stationarity-Ergodicity and for Comovements Between Nonlinear Discrete Time Markov Processes, *Journal of Econometrics*, 96, 39-73.
- ENDERS W. et GRANGER C. W. J. [1998], Unit-root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates, *Journal of Business and Economic Statistics*, 16(3), 304-11.
- et SIKLOS P. L. [2001], Cointegration and Threshold Adjustment, *Journal of Business and Economic Statistics*, 19(2), 166-176.
- ENGLE R. F. et GRANGER C. W. J [1997], Co-integration and Error Correction : Représentation, Estimation and Testing, *Econometrica*, 55(2), 251-76.
- ESCRIBANO A. et MIRA S. [2002], Nonlinear Error Correction Models, *Journal of Time Series Analysis*.
- EVANS M. et LEWIS K. [1994], Do Stationary Risk Premia Explain it all ? Evidence From the Term Structure, *Journal of Monetary Economics*, 33, 285-318.
- GILBERT R. A. [1985], Operating Procedures for Conducting Monetary Policy, *Federal Reserve Bank of St. Louis Review*, February, p. 13-21.
- GRANGER C. W. [1995], Modelling Nonlinear Relationships Between Extended Memory Variables, *Econometrica*, 63, 265-279.
- et TERASVIRTA T. [1993], *Modelling Non-Linear Economic Relationships*, Oxford : Oxford University Press.
- HANSEN B. et SEO B. [2002], Testing for Two-Regime Threshold Cointegration in Vector Error-Correction Models, *Journal of Econometrics*, 110, 293-318.
- HANSEN B. E. [1996], Inference when a Nuisance Parameter Is Not Identified Under the Null Hypothesis, *Econometrica*, 64(2), 413-430.
- JOHANSEN S. [1996], *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford, UK : Oxford University Press, 2nd edn.
- KAPETANIOS G. et SHIN Y. [2002], *Unit Root Tests in Three-Regime SETAR Models*, Working Paper n° 465, Department of Economics, University of London.
- , — et SNELL A., Testing for a unit root in the nonlinear STAR framework, *Journal of Econometrics*, 112(2), 359-379.
- MANKIW G. [1985], Small Menu Costs and Large Business Cycle : a Macro Model of Monopoly, *Quarterly Journal of Economics*, 100, 529-37.
- OBSTFELD M. et TAYLOR A. [1997], Nonlinear Aspects of Goods-Market Arbitrage and Adjustment : Heckscher's Commodity Points Revisited, *Journal of the Japanese and International Economies*, 11, 441-479.

PIPPENGER M. et GOERING G. [1993], A Note on the Empirical Power of Unit Root Tests Under Threshold Processes, *Oxford Bulletin of Economics and Statistics*, 55(4), 473-481.

— et — [2002], Additional results on the Power of Unit root and Cointegration Tests under Threshold Process, *Applied Economics Letters*, 7(10), 641-644.

SAIKKONEN P. [2002], *Stability Results for Nonlinear Vector Autoregressions with an Application to a Nonlinear Error Correction Model*, Manuscript, University of Helsinki.

SERCU P., UPPAL R. et VAN HULLE C. [1995], The exchange rate in the presence of transaction costs : implications for tests of purchasing power parity, *The Journal of Finance*, 50(4), 1309-19.

STOCK J. et WATSON M. [1988], Testing for Common Trends, *Journal of the American Statistical Association*, 83, 1097-1107.

TAYLOR A. [2001], Potential Pitfalls for the PPP Puzzle ? Sampling and Specification Biases in Mean-Reversion Tests of the LOOP, *Econometrica*, 69, 473-498.

TJØSTHEIM D. [1990], Non-Linear Time Series and Markov Chains, *Advances in Applied Probability*, 22, 587-611.

TONG H. et LIM K. [1980], Threshold Autoregression, Limit Cycles and Cyclical Data, *Journal of the Royal Statistical Society – Series B*, 42(3), 245-292.

TSAY R. [1998], Testing and Modeling Multivariate Threshold Models, *Journal of the American Statistical Association*, 93(443), 1188-1202.

TZAVALLIS E. et WICKENS M. [1998], A Re-Examination of the Rational Expectations Hypothesis of the Term Structure : Reconciling the Evidence from Long-Run and Short-Run Tests, *International Journal of Finance and Economics*, 3, 229-239.

Annexe

Tableau 3. Tests de cointégration de Johansen

Période	k	H ₀	λ – max	trace
(T)	10	r ≤ 1 r = 0	6,36 51,63	6,36 57,99
Test β = (1, -1) : χ ² (1) = 3,35 (p – val = 0,07)				
(T1)	7	r ≤ 1 r = 0	1,08 41,69	1,08 42,77
Test β = (1, -1) : χ ² (1) = 2,68 (p – val = 0,10)				
(T2)	16	r ≤ 1 r = 0	1,91 7,38	1,91 9,28
Test β = (1, -1) : χ ² (1) = 5,41 (p – val = 0,02)				
(T3)	7	r ≤ 1 r = 0	3,63 33,81	3,63 37,47
Test β = (1, -1) : χ ² (1) = 0,89 (p – val = 0,34)				

n.b. : Les valeurs critiques à 1 % et 5 % sont de (6,51 ; 15,69) et (3,84 ; 11,44) pour le λ – max et (6,51 ; 16,31) et (3,84 ; 12,53) pour la trace.

Tableau 4. Estimation du modèle à seuil par sous-période

(T1)
$\Delta S_t = \Gamma_{11}(L) \Delta S_{t-1} + \Gamma_{12}(L) \Delta R_{t-1}^S - 0,12 s_t S_{t-1} - 0,01 (1 - s_t) S_{t-1}$ $F_{11} = 0,00 \quad F_{12} = 0,00 \quad (-5,86) \quad (-0,15)$
$\Delta R_t^S = \Gamma_{21}(L) \Delta S_{t-1} + \Gamma_{22}(L) \Delta R_{t-1}^S + 0,19 s_t S_{t-1} - 0,11 (1 - s_t) S_{t-1}$ $F_{21} = 0,04 \quad F_{22} = 0,01 \quad (3,93) \quad (-1,07)$
$k = 7 ; L = -6\,444,8 ; \text{seuil} = 0,14 ; Q_1(12) = 0,92 ; Q_2(12) = 0,79$
<p>TESTS DE NON CAUSALITÉ :</p> <p>$s_t = 1 : F(7, 1009) = 6,38 (p - \text{val} = 0,00)$</p> <p>$s_t = 0 : F(7, 1009) = 1,91 (p - \text{val} = 0,06)$</p>
<p>TEST DE LINÉARITÉ :</p> <p>SupLM = 21,27 (bootstrap p – val = 0,002)</p>

(T2)			
$\Delta S_t =$	$\Gamma_{11}(L) \Delta S_{t-1}$	$+ \Gamma_{12}(L) \Delta R_{t-1}^S$	$- 0,11 s_t S_{t-1} + 0,14 (1 - s_t) S_{t-1}$
	$F_{11} = 0,00$	$F_{12} = 0,00$	$(- 2,14) \quad (1,28)$
$\Delta R_t^S =$	$\Gamma_{21}(L) \Delta S_{t-1}$	$+ \Gamma_{22}(L) \Delta R_{t-1}^S$	$- 0,05 s_t S_{t-1} - 0,01 (1 - s_t) S_{t-1}$
	$F_{21} = 0,01$	$F_{22} = 0,00$	$(- 0,35) \quad (- 0,02)$
$k = 16 ; L = - 1 298,2 ; \text{seuil} = 0,09 ; Q_1(12) = 0,34 ; Q_2(12) = 0,26$			
TESTS DE NON CAUSALITÉ :			
$s_t = 1 : F(16, 108) = 2,22 (p - \text{val} = 0,01)$			
$s_t = 0 : F(16, 108) = 2,21 (p - \text{val} = 0,01)$			
TEST DE LINÉARITÉ :			
Sup LM = 8,08 (bootstrap p - val = 0,339)			
(T3)			
$\Delta S_t =$	$\Gamma_{11}(L) \Delta S_{t-1}$	$+ \Gamma_{12}(L) \Delta R_{t-1}^S$	$- 0,12 s_t S_{t-1} - 0,02 (1 - s_t) S_{t-1}$
	$F_{11} = 0,00$	$F_{12} = 0,00$	$(- 5,19) \quad (- 0,45)$
$\Delta R_t^S =$	$\Gamma_{21}(L) \Delta S_{t-1}$	$+ \Gamma_{22}(L) \Delta R_{t-1}^S$	$+ 0,20 s_t S_{t-1} - 0,15 (1 - s_t) S_{t-1}$
	$F_{21} = 0,35$	$F_{22} = 0,00$	$(3,63) \quad (- 1,37)$
$k = 7 ; L = - 5 318,1 ; \text{seuil} = 0,14 ; Q_1(12) = 0,82 ; Q_2(12) = 0,74$			
TESTS DE NON CAUSALITÉ :			
$s_t = 1 : F(7, 840) = 4,82 (p - \text{val} = 0,00)$			
$s_t = 0 : F(7, 840) = 1,17 (p - \text{val} = 0,32)$			
TEST DE LINÉARITÉ :			
Sup LM = 10,79 (bootstrap p - val = 0,126)			

Les statistiques de Student sont données entre parenthèses.

$F_{ij} = p - \text{value}$ de l'hypothèse nulle que tous les coefficients de Γ_{ij} sont égaux à 0.

$Q_i(h) = p - \text{value}$ des tests d'autocorrélation des résidus à l'ordre h dans l'équation i.