

La prise en compte de ruptures dans l'évolution des
variables économiques :
Les modèles à changement de régimes

Corinne Perraudin
Université Paris I
SAMOS-MATISSE et EUREQua *

Octobre 2002

Abstract

Ce article se propose de présenter les méthodes économétriques en séries temporelles qui permettent de prendre en compte des ruptures dans l'évolution des variables économiques. Nous présentons différents modèles à changement de régimes : les modèles dans lesquels le changement de régime se fait selon les valeurs prises par une variable observable, dont la transition entre les régimes est brutale (modèle TAR) ou dont la transition est lisse (modèle STAR), les modèles à changement de régimes markovien développé initialement par Hamilton (1989).

Mots clés: Modèles à changement de régime; TAR; STAR; Markovien.

*Address: SAMOS, 90 rue de Tolbiac, 75013 Paris (France). Email: Corinne.Perraudin@univ-paris1.fr. La partie sur les modèles TAR est issue d'un travail avec Mélika Ben Salem, et la partie sur les modèles markoviens est issue d'un travail avec Frédéric Karamé.

1 Introduction

De nombreuses études récentes s'intéressent à la présence de non-linéarités dans les séries économiques, et proposent des procédures de tests permettant de mettre en évidence cette caractéristique. La propriété de non-linéarité des séries économiques peut être justifiée par l'existence d'asymétries dans la dynamique des variables. Neftçi [1984] trouve en effet que le taux de chômage augmente plus vite qu'il ne diminue. Beaudry et Koop [1993] montrent que les chocs favorables ont un effet plus important et plus persistant que les chocs défavorables. Afin de rendre compte de ces éventuelles asymétries, il est nécessaire d'avoir recours à des modèles économétriques capables de générer des dynamiques différentes selon la phase du cycle. Un processus de propagation linéaire ne répond pas à cette attente, puisqu'il implique une symétrie tout au long du cycle.

L'objectif de cet article est de montrer comment des méthodes récentes en séries temporelles non-linéaires peuvent être mises en œuvre et quelle peut être leur utilité. Nous étudions plus particulièrement la classe des modèles à changement de régime. Ces modèles permettent en effet d'engendrer des asymétries, puisqu'ils sont définis par plusieurs régimes aux dynamiques distinctes. On peut donc espérer, avec de tels modèles, repérer des régularités distinctes dans les phases ascendantes et descendantes. Ces modèles permettent de plus de rendre compte de ruptures de grande ampleur dans l'évolution des variables ou d'ajustements non continus. Par exemple, la présence de coûts fixes conduit les agents à ne pas s'ajuster de manière continue. C'est seulement quand la déviation à l'équilibre atteint un certain seuil qu'il y a ajustement. Ce type de modélisation a été utilisé pour expliquer la dynamique des stocks, des prix ou de l'emploi¹. La présence d'ajustements discrets peut correspondre aussi aux interventions de politique économique. Par exemple, dans le cadre d'une zone de change, les taux de change peuvent varier à l'intérieur d'une bande, mais s'ils atteignent des valeurs correspondant aux bornes, les Banques centrales interviennent de manière ponctuelle sur le marché des changes². Ces divers exemples illustrent le regain d'intérêt pour la famille des modèles à seuils.

Bien que la littérature empirique dans ce domaine soit en pleine expansion, les travaux concernant la spécification de ces modèles n'en sont qu'à leurs débuts. Comme l'intérêt est autant de détecter des non-linéarités que de les caractériser et de les estimer, nous étudierons les tests de linéarité définis contre l'alternative donnée par les modèles à changement de régime.

Nous présentons trois types de modèles à changement de régimes, selon la manière de définir la variable qui implique un changement de régime et selon la nature de la transition entre les régimes. Parmi les modèles dans lesquels le changement de régime

¹Voir par exemple Perraudin [1995], Perraudin [1999], Karamé et Perraudin [1999] et BenSalem et Perraudin [2001].

²Voir par exemple Bec, Ben Salem et MacDonald [1999] et Bec, Ben Salem et Collard [2002].

se fait selon les valeurs prises par une variable observable, on présentera successivement le cas où la transition est brutale (modèle TAR) puis le cas où la transition est lisse (modèle STAR). Ensuite seront présentés les modèles à changement de régimes markovien développé initialement par Hamilton (1989).

2 Modèles à changement de régime brutal: les modèles TAR

Avant d'estimer un modèle à changement de régime brutal, des auteurs ont proposés des tests de linéarité contre l'alternative fournie par ces modèles afin de s'assurer que cette représentation est pertinente pour les données dont on dispose. Nous présentons ici deux méthodes alternatives, celle de Tsay (1989) et celle de Hansen (1996). Les deux méthodes diffèrent par le traitement des paramètres de nuisance dans la construction des statistiques de tests.

La première section présente les différentes représentations des modèles à seuil, et discute des difficultés associées à la spécification de ces modèles. La procédure de Tsay [1989] – tests de linéarité et détection du seuil définissant les deux régimes – est exposée dans la deuxième section. La troisième section présente la procédure de Hansen [1996a].

2.1 Les modèles à seuil

L'existence de relations non-linéaires entre les variables économiques peut être modélisée de façon simple à l'aide des modèles à seuil. En effet, la linéarisation par morceaux est souvent une approximation correcte de la dynamique non-linéaire d'un processus. Ainsi, dans les modèles à seuil, cette dernière est représentée par des dynamiques linéaires différentes selon la situation du système. Le changement de dynamique s'opère selon le signe de l'écart entre la variable de transition et la valeur du seuil.

2.1.1 Définition des modèles à seuil

Les modèles à seuil ont été introduits par Tong et Lim [1980] (voir aussi Tong [1990]). Les applications mettent en valeur l'existence de dynamiques asymétriques au cours du cycle. Sur le taux de croissance du PIB américain, Potter [1995] montre que le modèle autorégressif standard est rejeté au profit d'un modèle à seuil³. Il montre de plus que ce modèle rend correctement compte de l'asymétrie cyclique existante, en étudiant des fonctions de réponse non-linéaires suite à des chocs de taille et de signe différents.

Le processus Y_t , supposé stationnaire⁴, suit un modèle à seuil à deux régimes défini

³Il impose un seuil égal à zéro et la variable de transition est l'endogène retardée de deux périodes.

⁴Il faut noter que l'inférence dans ces modèles requiert la stationnarité de la série étudiée. Les études empiriques modélisent alors le taux de croissance des séries. Néanmoins, des travaux récents suggèrent que la présence d'une racine unitaire peut être liée à une mauvaise spécification du processus, notamment linéaire (Caner et Hansen [1997]). A l'instar des études empiriques s'attachant à la détection de non-linéarités, nous étudierons les taux de croissance des variables, puisque notre objectif est de

de la façon suivante :

$$Y_t = \begin{cases} \phi_0^{(1)} + \phi^{(1)} X_t + \epsilon_t^{(1)} & \text{si } Z_t \leq s \\ \phi_0^{(2)} + \phi^{(2)} X_t + \epsilon_t^{(2)} & \text{si } Z_t > s \end{cases} \quad (1)$$

où $\epsilon_t^{(j)}, j = 1, 2 \quad t = 1, \dots, n$, deux bruits blancs indépendants de variance $\sigma_{(j)}^2$, n étant le nombre d'observations, $X_t = (Y_{t-1}, \dots, Y_{t-p}, V_1, \dots, V_k)'$ et $\phi^{(j)} = (\phi_1^{(j)}, \dots, \phi_m^{(j)})$, $m = p + k$. Les $V_i, i = 1, \dots, k$, sont des variables explicatives. Ce modèle correspond à une régression non-linéaire. Quand $X_t = (Y_{t-1}, \dots, Y_{t-p})'$, le modèle est alors autorégressif par morceaux, et on parle de modèle TAR, pour *Threshold AutoRegressive*.

La variable de transition, Z_t , est une des variables de X_t , c'est-à-dire une endogène retardée, Y_{t-d} (et d est appelé le paramètre de délai), ou une variable explicative⁵. Le choix de la variable de transition peut être guidé par la théorie économique, mais une solution athéorique est aussi envisageable. Comme nous le verrons ensuite, la procédure de spécification du modèle à seuil permet alors de décider quelle variable retenir.

Le paramètre s est le seuil, pour lequel il y a changement d'état du système. Dans le cadre d'études conjoncturelles, une valeur du seuil égale à zéro définit deux régimes : un régime de croissance positive et un régime de croissance négative. La valeur du seuil fournit une première interprétation économique des régimes définissant la dynamique du processus. Il est donc important de trouver une bonne méthode d'estimation ou de détection de ce seuil.

Lorsque les observations de la variable de transition ont une valeur inférieure au seuil s , la dynamique de Y est caractérisée par les paramètres $\phi_i^{(1)}, i = 0, \dots, m$. Quand la variable de transition prend des valeurs supérieures au seuil, la dynamique de Y est expliquée par les coefficients $\phi_i^{(2)}, i = 0, \dots, m$. Ainsi, la dynamique de Y est linéaire par morceaux, et permet de rendre compte de comportements différents selon les valeurs prises par la variable Z . Le modèle dépend ainsi d'une manière non-linéaire du paramètre définissant la variable de transition (le délai) et du seuil. Il est possible d'étendre cette spécification à un modèle à plus de deux régimes en définissant alors plusieurs seuils.

Une écriture équivalente du modèle à seuil (1) est obtenue en introduisant une fonction de transition :

$$Y_t = (\phi_0^{(1)} + \phi^{(1)} X_t) \mathbb{I}\{Z_t \leq s\} + (\phi_0^{(2)} + \phi^{(2)} X_t) \mathbb{I}\{Z_t > s\} + \epsilon_t \quad (2)$$

où ϵ_t est une séquence de bruits blancs i.i.d. qui vaut $\epsilon_t^{(1)} \mathbb{I}\{Z_t \leq s\} + \epsilon_t^{(2)} \mathbb{I}\{Z_t > s\}$, et de variance $\sigma_\epsilon^2 = \sigma_{(1)}^2 \mathbb{I}\{Z_t \leq s\} + \sigma_{(2)}^2 \mathbb{I}\{Z_t > s\}$.

comparer les performances des modèles linéaires et non-linéaires.

⁵A la différence de ce que propose Tong [1990], Caner et Hansen [1997] ne supposent pas que la variable de transition est une des variables explicatives. Or, puisque cette variable détermine le changement de régimes, il est naturel de l'introduire parmi les explicatives.

$\mathbb{I}\{\cdot\}$ est une fonction indicatrice qui vaut 1 quand l'inégalité est vérifiée et 0 sinon. Il existe donc deux régimes distincts et la transition entre ces deux régimes est brutale. Ceci est un cas particulier des modèles à seuil à transition lisse proposés par la suite par Teräsvirta [1994], sous le vocable de modèles STAR pour *Smooth Transition AutoRegressive*, que nous présenterons dans la section suivante.

Avant de présenter les procédures de spécification des modèles à seuil à transition brutale proposées par Tsay [1989] et par Hansen [1996a], quelques stratégies communes aux deux méthodes de spécification sont d'abord énoncées.

2.1.2 Spécification du modèle à seuil

La spécification du modèle à seuil nécessite de choisir certains paramètres, qui ne peuvent être estimés par des méthodes usuelles. Il s'agit du nombre de régresseurs ainsi que du paramètre définissant la variable de transition. Ils sont alors sélectionnés préalablement à l'estimation des autres paramètres du modèle. Les tests de linéarité contre l'alternative donnée par le modèle à seuil guident ces choix.

Tests de linéarité

De nombreux tests de linéarité existent dans la littérature⁶. Une première possibilité est de tester la linéarité sans spécifier d'alternative précise. Le plus connu est peut-être le test RESET, pour *Regression Error Specification Test* proposé par Ramsey [1969] (voir aussi Keenan [1985]). L'idée est de tester un modèle linéaire contre un modèle augmenté de termes élevés à une puissance supérieure à deux. Ce test consiste alors à tester la nullité de ces derniers termes. L'inconvénient des tests sans alternative spécifiée se révèle dans le cas où la linéarité est rejetée.

Une seconde possibilité consiste à tester un modèle linéaire contre un modèle non-linéaire donné. En effet, pour la plupart des modèles non-linéaires envisagés dans la littérature, on peut définir des contraintes pour lesquelles le modèle devient linéaire. L'hypothèse de linéarité revient alors à tester cet ensemble de contraintes. Dans ce cas, si la linéarité est rejetée, on connaît l'alternative à considérer. Pour ces tests, la présence éventuelle de paramètres de nuisance non identifiables sous l'hypothèse nulle de linéarité conduit souvent à utiliser une approximation du modèle sous l'hypothèse alternative (par exemple, en utilisant un développement limité de la partie non-linéaire). Cette stratégie peut conduire à une situation où le test est puissant contre différentes alternatives⁷. L'estimation du modèle non-linéaire permet alors de conclure quant au choix de l'alternative. Il est important de noter que le non-rejet de la linéarité ne

⁶Voir Tong [1990], Granger et Teräsvirta [1993], Guégan [1994] pour une revue de la littérature sur les divers modèles non-linéaires et les tests de linéarité.

⁷Dans le cas d'une alternative de modèles à seuil à transition lisse, le test de linéarité proposé par Luukkonen, Saikkonen et Teräsvirta [1988] repose sur le développement de Taylor de la fonction de transition, qui identifie aussi bien un modèle bilinéaire qu'un modèle ARCH. Le test est puissant contre ces deux alternatives particulières en plus du modèle à seuil à transition lisse.

signifie pas forcément que le processus est linéaire, l'alternative peut aussi être mal spécifiée.

Nous choisissons de tester l'hypothèse de linéarité contre l'alternative fournie par les modèles à seuil. Il s'agit alors de tester l'hypothèse nulle suivante :

$$H_0 : \phi_0^{(1)} = \phi_0^{(2)} \text{ et } \phi_i^{(1)} = \phi_i^{(2)} \quad i = 1, \dots, m$$

contre l'alternative fournie par le modèle à seuil (1). Ainsi, le test requiert la spécification préalable du modèle linéaire, *i.e.* le choix des m régresseurs.

Il faut noter que, sous l'hypothèse nulle de linéarité, le paramètre de délai — ou la variable de transition — et le seuil ne sont pas identifiables. En effet, la valeur de la vraisemblance du modèle linéaire ne dépend pas de ces paramètres si on suppose que la variance des deux bruits blancs est la même dans les deux régimes. Ainsi, ils peuvent prendre n'importe quelles valeurs sans changer la fonction de vraisemblance. Les tests usuels d'un modèle contre un autre (statistiques de Wald, du ratio de vraisemblance ou du multiplicateur de Lagrange) ne sont alors pas applicables. Face à ce problème, Tsay [1989] et Hansen [1996a] proposent deux méthodes distinctes, comme nous le verrons lors de la présentation de ces procédures. Elles diffèrent notamment quant au traitement du seuil, la variable de transition étant traitée de la même manière. En effet, les deux procédures de test sont développées pour une variable de transition choisie préalablement. Les paragraphes suivants exposent le choix des régresseurs et de la variable de transition.

Choix des régresseurs

Le choix des régresseurs, *i.e.* les k variables explicatives et les p retards de l'endogène, du modèle à seuil s'effectue dans le cadre d'un modèle linéaire de référence. Les régresseurs sont sélectionnés sur la base de critères d'information (AIC, BIC, Hannan), du test du portemanteau et/ou du test de significativité du retard le plus élevé présent dans la régression. Cette première étape est très importante puisque le modèle linéaire constitue l'hypothèse nulle du test de linéarité. Comme le fait remarquer Teräsvirta [1994] dans le cadre d'un modèle STAR, si le retard autorégressif est surestimé, cela peut avoir un effet pervers sur la puissance du test (comparé au cas où p serait connu). Alternativement, si le retard sélectionné est trop faible et qu'il reste de l'autocorrélation résiduelle, les tests de linéarité sont biaisés (la linéarité est rejetée à tort).

2.1.3 Choix de la variable de transition

Le choix de la variable de transition, ou l'estimation du paramètre de délai dans un modèle TAR, ne relève pas des méthodes usuelles. En effet, la fonction de vraisemblance n'est pas dérivable par rapport à ce paramètre. Il faut avoir recours à d'autres méthodes.

La variable de transition est alors choisie par une procédure particulière, basée sur les résultats des tests de linéarité. L'idée est que la linéarité sera d'autant plus fortement rejetée que le modèle à seuil est bien spécifié. Ainsi les tests de linéarité sont menés pour une variable de transition donnée. Cependant, celle-ci n'étant pas connue, on construit le test de linéarité pour toutes les variables de transition possibles, *i.e.* pour l'ensemble des variables X_t ou pour toutes les valeurs du paramètre de délai d telles que $1 \leq d \leq p$ dans le cas du modèle TAR. On retient ensuite la variable de transition pour laquelle la linéarité est la plus fortement rejetée, c'est-à-dire celle qui minimise la probabilité du test. En effet, le test de linéarité est d'autant plus puissant que la variable de transition est bien choisie.

Estimation

Pour un ensemble donné de variables explicatives, et pour une valeur du paramètre de délai et du seuil⁸, les estimateurs par moindres carrés ordinaires des coefficients autorégressifs $\phi_i^{(j)}$, $i = 0, \dots, m$, du régime j sont asymptotiquement normaux (voir Chan et Tong [1986]). Ils minimisent la somme des carrés des résidus du régime j :

$$SCR^{(j)} = \sum_{t \in K(j)} \left(Y_t - \phi_0^{(j)} - \phi^{(j)} X_t \right)^2$$

où $K(j)$ est l'ensemble $\{p < t \leq n : Z_t \text{ appartient au régime } j\}$. L'estimateur de la matrice de variance-covariance des coefficients estimés est donné par :

$$\widehat{V}^{(j)} = \hat{\sigma}_{(j)}^2 (\mathcal{X}' \mathcal{X})^{-1}$$

où $\mathcal{X} = (1, X_t)_{t \in K(j)}$, $\hat{\sigma}_{(j)}^2 = \frac{SCR^{(j)}}{n_j - (m+1)}$ et n_j est le nombre d'observations dans le régime j .

Les méthodes de spécification du modèle à seuil proposées par Tsay [1989] et par Hansen [1996] incorporent un test de linéarité et une procédure de détection du seuil. Leur méthode respective est présentée afin d'en souligner les différences et les similitudes. Cette comparaison se poursuit par l'examen des performances de chaque test de linéarité à travers des exercices de puissance.

2.2 La méthodologie de Tsay [1989]

Cette section présente la méthodologie proposée par Tsay, pour tester l'hypothèse de linéarité contre l'alternative fournie par le modèle à seuil, ainsi que la procédure graphique de détection du seuil.

⁸Le seuil est déterminé de manière différente selon la procédure de spécification retenue, comme cela sera présenté dans la section suivante.

2.2.1 Test de linéarité

Tsay [1989] ne traite pas le paramètre de seuil de la même manière que le paramètre définissant la variable de transition. Le paramètre de délai est fixé lors de la mise en œuvre du test de linéarité, et comme il a été mentionné précédemment, on retient la variable de transition pour laquelle la linéarité est la plus fortement rejetée.

L'originalité de la méthode de Tsay [1989] est de se débarrasser du paramètre de nuisance (le seuil), dans la logique des tests CUSUM, adaptés aux modèles à seuil par Petrucelli et Davies [1986]. Tsay réécrit l'alternative du test de linéarité en réarrangeant toutes les observations selon l'ordre croissant de la variable de transition (voir encadré 1). La régression ordonnée ainsi obtenue est constituée d'une première régression linéaire pour les r observations correspondant aux valeurs de la variable de transition inférieures au seuil. La deuxième régression correspond à l'autre (ou aux autres si le modèle (1) est constitué de plus de deux régimes) régime.

La régression ordonnée (3) est alors estimée de façon récursive (voir Ertel et Fowlkes [1976] et Tsay [1989]). L'intérêt de cette écriture (et de l'estimation récursive) est qu'elle ne nécessite pas la connaissance du seuil ni du nombre d'observations dans le premier régime (soit r).

En effet, si le processus sous-jacent aux données est linéaire — toutes les observations appartiennent au même régime — les résidus récursifs ainsi obtenus sont des bruits blancs asymptotiquement et sont orthogonaux aux régresseurs. En revanche, s'il existe deux régimes distincts, la suite de résidus récursifs à partir de l'observation correspondant au seuil n'est plus orthogonale aux régresseurs. Elle devient fonction des régresseurs. Dans ce cas, la projection de la série de résidus récursifs sur les régresseurs conduit à des coefficients estimés significativement différents de zéro. Le test de linéarité revient à tester la nullité de ces coefficients. La statistique de test est alors indépendante du paramètre de nuisance s , et suit un Fisher standard. Elle dépend par contre du paramètre de délai et du nombre de régresseurs m que l'on suppose fixés pendant le test.

Encadré 1 : Procédure de spécification de Tsay [1989]

- (i) Choix des régresseurs $X_t = (Y_{t-1}, \dots, Y_{t-p}, V_1, \dots, V_k)'$ dans le cadre d'un modèle linéaire :

$$Y_t = \phi_0 + \phi X_t + \epsilon_t, \quad \epsilon_t \text{ i.i.d.}(0, \sigma^2)$$

- (ii) Pour chaque variable de transition possible, soit $Z_t = X_{it}$, $i = 1, \dots, m$ ($m = p + k$) :

- Estimation des régressions ordonnées de manière récursive :

$$Y_{(l)} = \begin{cases} \phi_0^{(1)} + \phi^{(1)} X_{(l)} + \epsilon_{(l)}^{(1)} & \text{pour les } r \text{ 1}^{\text{ères}} \text{ observations de } Z_L \\ \phi_0^{(2)} + \phi^{(2)} X_{(l)} + \epsilon_{(l)}^{(2)} & \text{ensuite} \end{cases} \quad (3)$$

où Z est la variable de transition et L l'ordre correspondant au classement des observations de la variable de transition dans le sens croissant tel que $Z_{(l)} \leq Z_{(l+1)}$ ($(l) \in L$). Cette régression ordonnée en (l) correspond à la réécriture du modèle (1) quand le seuil se situe entre la $r^{\text{ème}}$ et la $(r+1)^{\text{ème}}$ observation ordonnée.

L'hypothèse nulle de linéarité devient : $H'_0 : \phi_i^{(1)} = \phi_i^{(2)}$ ($i = 0, \dots, m$) dans le modèle (3).

- Construction de la statistique de test de linéarité $Q(m)$ et calcul de la probabilité du test associée :

$$Q(m) = \frac{(\sum \hat{\epsilon}_t^2 - \sum \hat{a}_t^2)/(m+1)}{\sum \hat{a}_t^2/(n-b-2m-1)}$$

où b est le nombre d'observations utilisées pour initialiser les estimations récursives, $\hat{\epsilon}_t$ sont les résidus ϵ_t standardisés et \hat{a}_t sont les résidus de la régression de $\hat{\epsilon}_{(l)}$ sur $(1, X'_{(l)})$. Sous l'hypothèse nulle de linéarité, la statistique $Q(m)$ suit un Fisher à $(m+1)$ et $(n-b-2m-1)$ degrés de liberté. Comme le conseille Tsay, on peut prendre $b = \frac{n}{10} + p$ dans le cas d'un modèle autorégressif de type TAR. Dans le cas d'une régression, nous proposons $b = \text{Max}(\frac{n}{10}, m+3)$.

- (iii) Choix de la variable de transition : celle qui maximise la statistique $Q(m)$ sachant que la linéarité est rejetée.
- (iv) Détection du seuil par analyse graphique des coefficients estimés (de manière récursive) et de leur Student en fonction de la variable de transition ordonnée (retenue dans l'étape précédente).
- (v) Estimation du modèle à seuil, sachant m , la variable de transition et le seuil.

2.2.2 La détection du seuil

L'estimation récursive des régressions ordonnées permet d'estimer des résidus récursifs, et de calculer un certain nombre de statistiques récursives (coefficients et Student) utilisées pour la localisation du seuil. En effet, dans le cadre de ces régressions ordonnées, le processus associé aux r premières observations (ordonnées) est linéaire. Dès que l'on introduit l'observation dont la valeur correspond au seuil, le processus correspond alors à une combinaison de deux (ou plus) régressions. Ainsi le profil des résidus récursifs et des autres statistiques sera déformé au niveau du seuil. En effet, ils dépendent alors des régresseurs $(1, X'_{(l)})$. La localisation du seuil se fait donc par analyse graphique des représentations des différentes statistiques récursives en fonction de la variable de transition ordonnée.

Prenons l'exemple de la statistique de Student afin de donner l'idée de cette procédure : tant que les observations sont inférieures à la valeur du seuil, le modèle est linéaire. De plus, si un coefficient est significatif, son Student converge progressivement vers une valeur au fur et à mesure qu'on ajoute des observations. Quand le seuil est atteint, les estimations changent et la statistique de Student aussi (elle change parfois de direction pour la valeur du seuil). Tsay fait remarquer que les représentations des statistiques de Student sont intéressantes, puisqu'elles indiquent, en plus des changements de direction ou des sauts, la significativité des variables. De plus, le Student de la constante est important dans la mesure où il révèle les changements dans le niveau. L'analyse du coefficient du régresseur correspondant à la variable de transition permet, en plus de localiser le seuil, de repérer la nature de la transition : lisse ou brutale, logistique ou exponentielle.

La détection des seuils consiste alors à localiser les ruptures dans les différentes représentations graphiques. Seule la première rupture est pertinente puisqu'au delà de ce seuil, la régression ordonnée peut être une combinaison de plusieurs régimes. Ainsi, une fois le seuil détecté, il est alors possible de réestimer les régressions ordonnées sur les observations pour lesquelles la variable de transition est supérieure au seuil détecté, afin de vérifier s'il n'existe pas d'autres régimes. On retient pour valeur du seuil celle qui minimise le critère AIC (des modèles à seuil) au voisinage de la valeur du seuil précédemment localisé.

En présence de points aberrants, ou *outliers*, suffisamment importants, les statistiques récursives peuvent être biaisées. Cependant seuls les points aberrants de signe négatif affectent la procédure de sélection des seuils.

Si la valeur du seuil correspond à une observation qui a servi à l'initialisation des estimations récursives, la méthode de Tsay ne permettra pas d'identifier ce seuil. Ce problème peut arriver aussi si la valeur du seuil est très élevée, de telle sorte que le nombre d'observations intégrées avant la rupture est très grand. Une possibilité est alors de réordonner les observations, non pas dans le sens croissant de la variable de transition, mais dans le sens décroissant. Cependant, Tsay (voir aussi Petrucelli et Davies [1986], Petrucelli [1990]) suggère uniquement le réarrangement des observations selon l'ordre croissant de la variable de transition.

Même si cette méthode n'est qu'une procédure graphique, elle fournit des indications utiles sur la localisation des seuils et sur le nombre de régimes du modèle à seuil.

2.3 La méthodologie de Hansen [1996a]

Cette section présente la méthodologie proposée par Hansen, qui permet de tester l'hypothèse de linéarité contre l'alternative fournie par le modèle à seuil, ainsi que la procédure de sélection du seuil.

2.3.1 Tests de linéarité

Face à la présence de paramètres de nuisance non identifiés sous l'hypothèse nulle de linéarité, Hansen [1996a] construit le test de linéarité en considérant le paramètre qui définit la variable de transition (paramètre de délai dans le cas du modèle TAR) fixé, comme le fait Tsay. Mais contrairement à Tsay, il traite de la même façon le paramètre de seuil et le paramètre définissant la variable de transition. Ainsi, il ne réécrit pas l'alternative à la linéarité comme Tsay, mais construit un test qui dépend de ces deux paramètres.

La statistique de test Il choisit d'utiliser des statistiques de tests usuelles, telles que LM (*Lagrange Multiplier*), Wald et LR (*Likelihood Ratio*), qui vont dépendre de s et de d . Comme ces paramètres sont choisis, non pas *a priori* c'est-à-dire indépendamment des observations, mais en vue de maximiser la fonction de vraisemblance, le test et sa distribution ne sont pas connus. Parmi les trois statistiques de tests, LM est retenue dans la mesure où elle ne requiert que l'estimation du modèle sous H_0 , donc du modèle linéaire.

Pour une variable de transition donnée, on calcule alors une séquence de statistiques $LM(s)$ à partir d'un balayage sur toutes les valeurs possibles du paramètre de seuil s , c'est-à-dire sur toutes les observations de la variable de transition⁹.

Le test le plus puissant contre l'alternative de modèles à seuil correspond à la statistique la plus élevée. Ainsi, à partir de cette séquence de statistiques $LM(s)$, Davies [1977] préconise de prendre $Sup_s LM(s)$, puisque la statistique LM est une fonction monotone du paramètre de seuil. Andrews et Ploberger [1994] montrent que les tests optimaux en présence de paramètres de nuisance non identifiés sous l'hypothèse nulle ont une forme moyenne exponentielle. Les statistiques de test de linéarité utilisées sont : la statistique $SupLM$, ainsi que la moyenne de la séquence de statistiques $LM(s)$ obtenue en intégrant par rapport à s (notée ensuite $AveLM$) et le logarithme de la moyenne de la séquence des statistiques $LM(s)$ transformées de la façon suivante : $exp(\frac{1}{2}LM(s))$ (notée ensuite $ExpLM$). La variable de transition retenue est alors celle qui maximise les statistiques de tests $SupLM$, $AveLM$ et $ExpLM$.

La distribution asymptotique La distribution asymptotique de ces statistiques n'étant pas connue, Hansen propose de l'approximer par une méthode de bootstrap. Ceci consiste à calculer un point de la distribution asymptotique à partir d'un tirage particulier dans l'ensemble des résidus du modèle. En répliquant cette procédure, on obtient l'estimation de la distribution asymptotique et, à défaut des valeurs critiques qui ne peuvent être tabulées, les probabilités de test correspondantes. On obtient ces

⁹Nous retirons les 15% premières et dernières valeurs des observations classées afin que chaque régime contiennent un nombre suffisant d'observations.

probabilités en calculant, parmi les J statistiques de test simulées¹⁰, la proportion de celles qui excèdent la valeur estimée de la statistique considérée.

Encadré 2 : Procédure de spécification de Hansen [1996a]

- (i) Choix des régresseurs $X_t = (Y_{t-1}, \dots, Y_{t-p}, V_1, \dots, V_k)'$ dans le cadre d'un modèle linéaire :

$$Y_t = \phi_0 + \phi X_t + \epsilon_t$$

- (ii) Pour chaque variable de transition possible, soit $Z_t = X_{it}$, $i = 1, \dots, m$ ($m = p + k$) :

- Construction de la séquence de statistiques de test de linéarité $LM(s)$ pour toutes les valeurs possibles du seuil s , c'est-à-dire pour toutes les observations de la variable de transition:

$$LM(s) = S_n(s)' M_n(s) S_n(s)$$

avec $S_n(s) = \frac{1}{n} \sum_{t=1}^n X_t \hat{\epsilon}_t(s)$ le score du modèle à seuil estimé sous H_0 . $M_n(s)$ est la matrice d'information de Fisher, estimée sous H_0 , corrigée ou non de l'hétéroscédasticité.

- Calcul des statistiques $SupLM$, $ExpLM$ et $AveLM$ et de leur probabilité du test associée (par bootstrap):

$$\hat{p}_n = \frac{1}{J} \sum_{j=1}^J \mathbb{I} \left\{ g(LM(s)) \leq g(LM(s))^j \right\}$$

où $g(\cdot) = \{Sup, Ave, Exp\}$.

- (iii) Choix de la variable de transition : celle qui maximise les statistiques $SupLM$, $ExpLM$ et $AveLM$ sachant que la linéarité est rejetée.
- (iv) Estimation du seuil qui minimise la variance résiduelle du modèle à seuil estimé parmi toutes les valeurs possibles de la variable de transition.
- (v) Estimation du modèle à seuil, sachant m , la variable de transition et le seuil.
-

¹⁰ J est le nombre de simulations. On peut choisir $J = 1000$, qui est un bon compromis entre un nombre important de simulations pour que l'estimation des probabilités de test soit suffisamment précise et un nombre de simulations raisonnable quant au temps de calcul.

2.3.2 Estimation du seuil

Hansen [1996b] préconise d'utiliser la méthode des moindres carrés séquentiels pour estimer le seuil.

Pour une valeur donnée du seuil s , les estimateurs des paramètres $\phi^{(j)}$, $j = 1, 2$ sont obtenus par moindres carrés par régime. On évalue alors la variance résiduelle du modèle à seuil conditionnellement à cette valeur du seuil :

$$\hat{\sigma}_n^2(s) = \frac{1}{n} \sum_{t=1}^n (\hat{\epsilon}_t(s))^2$$

où $\hat{\epsilon}_t(s) = Y_t - (\phi_0^{(1)} + \phi^{(1)} X_t) \mathbb{I}\{Z_t \leq s\} - (\phi_0^{(2)} + \phi^{(2)} X_t) \mathbb{I}\{Z_t > s\}$.

Ces estimations sont effectuées pour toutes les valeurs possibles du seuil, c'est-à-dire pour toutes les observations de la variable de transition. L'estimation du seuil est alors la valeur qui minimise la variance résiduelle :

$$\hat{s} = \text{ArgMin}_s \hat{\sigma}_n^2(s)$$

Hansen [1996b] propose un intervalle de confiance pour le paramètre s , basé sur la distribution asymptotique de la statistique du rapport de vraisemblance. L'intervalle de confiance est alors donné par :

$$\hat{\Gamma} = \{s : \hat{\sigma}_n^2(s) \leq \sigma_n^2(1 + c_\xi(\beta)/n)\}$$

où $c_\xi(\beta)$ est la valeur critique au niveau β tabulée par Hansen [1996b].

BenSalem et Perraudin [2001] présentent quelques exercices de simulation permettant de comparer les performances en matière de taille et de puissance des deux tests de linéarité, et d'étudier la puissance de la procédure d'estimation du seuil proposée par Hansen.

Ces exercices de puissance révèlent l'importance du choix de p et l'effet de la correction de l'hétéroscédasticité. La surestimation de p ou une mauvaise estimation du délai de la variable de transition sont très coûteuses en termes de puissance, tandis que la correction de l'hétéroscédasticité altère la puissance des tests.

Les auteurs indiquent aussi que la taille des tests de Tsay et de Hansen est comparable, mais que la correction de l'hétéroscédasticité altère la taille, dans le sens d'un rejet trop fréquent de l'hypothèse nulle, tout particulièrement pour SupLM.

Concernant la procédure d'estimation du seuil proposée par Hansen¹¹, il apparaît que le seuil est estimé de plus en plus précisément quand le modèle s'éloigne de l'hypothèse de linéarité. Cependant, dans le cas d'un changement dans le paramètre autorégressif, il apparaît que la procédure de sélection du seuil est peu satisfaisante,

¹¹Il est difficile de proposer une approximation de la distribution de l'estimateur du seuil dans la procédure de spécification de Tsay, celle-ci étant une méthode graphique donc informelle.

puisque la distribution obtenue est très étalée. Ainsi, les auteurs concluent qu'il ne faut pas se contenter de la méthode de minimisation de la variance résiduelle proposée par Hansen pour obtenir une estimation du seuil. La confrontation avec les résultats de la méthode graphique de Tsay paraît indispensable.

3 Les modèles à changement de régime lisse

Cette section présente les modèles à seuil, dans lesquels la transition entre les régimes est continue et plus ou moins rapide. Ceci permet alors d'expliquer des comportements de lissage des choix. L'avantage de ces modèles est d'engendrer des dynamiques moins extrêmes que dans le cadre des modèles TAR présentés dans la section précédente, puisqu'il existe alors un continuum de régimes entre les deux dynamiques extrêmes. De plus, la méthodologie de spécification des modèles à transition lisse permet de choisir entre différents types de transition, mais aussi d'estimer la vitesse de transition, plutôt que de l'imposer infinie. Enfin, le seuil est estimé comme les autres paramètres autorégressifs, et non pas déterminé dans la première étape de la spécification. La variable de transition est en revanche toujours déterminée lors de la première étape.

3.1 Présentation des modèles à seuil à transition lisse

Chan et Tong [1986] ont suggéré de généraliser le modèle TAR en permettant à la transition entre les régimes d'être lisse. La spécification et l'estimation de ces modèles, appelés modèles STAR pour *Smooth Transition AutoRegressive* ont été développées par Teräsvirta [1994]¹². Ces modèles semblent adéquats pour la modélisation d'indicateurs du cycle économique¹³. En effet, ils sont composés de deux régimes extrêmes, mais de par la nature de la transition, le passage entre ces régimes n'est pas brutal. Il existe un continuum de régimes intermédiaires. De plus, la nature de la transition est donnée par les résultats des tests de linéarité et de spécification du modèle. Elle n'est donc pas imposée *a priori*. Il en est de même de la vitesse de transition entre les régimes. Les modèles STAR peuvent engendrer de l'asymétrie puisque la dynamique des régimes alternatifs est différente. Cette section présente ces modèles, ainsi que leur procédure de spécification et d'estimation.

3.1.1 Définition du modèle STAR

On définit le modèle STAR(p) de la façon suivante :

$$Y_t = \left(\phi_0^{(1)} + \phi^{(1)'} X_t \right) (1 - F(Z_t; \gamma, s)) + \left(\phi_0^{(2)} + \phi^{(2)'} X_t \right) F(Z_t; \gamma, s) + \epsilon_t \quad (4)$$

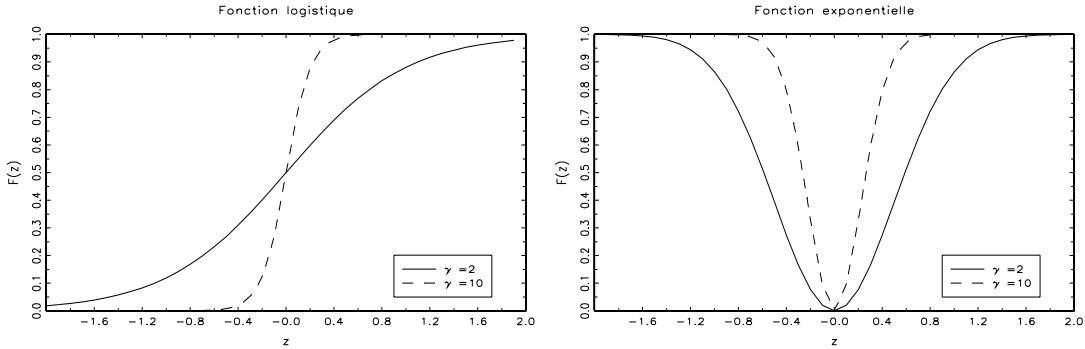
¹²Voir aussi les travaux de Luukkonen, Saikkonen et Teräsvirta [1988a], Teräsvirta et Anderson [1993].

¹³Même si ce modèle semble plus pertinent pour rendre compte des fluctuations conjoncturelles, Teräsvirta [1995] utilise cette représentation pour décrire la réponse du taux de croissance annuel américain entre 1889 et 1987 à des chocs exogènes importants. Il montre qu'elle peut être utilisée aussi pour expliquer des changements brusques et soudains.

avec $\epsilon_t \sim N(0, \sigma^2)$, $\gamma > 0$. Les autres notations sont les mêmes que celles données dans la section précédente pour les modèles TAR.

F est la fonction de transition. C'est une fonction comprise entre 0 et 1. Elle dépend de la variable de transition Z_t (Y_{t-d} dans le cas univarié), du seuil s , et du paramètre de lissage γ , qui gouverne la vitesse de transition entre les régimes. Plus il est élevé et plus la transition entre les deux régimes est brutale. La fonction de transition F peut prendre différentes formes. Si c'est une fonction dichotomique qui prend les valeurs 0 ou 1, on retrouve un modèle à seuil à transition brutale (TAR) étudié dans la section précédente. Teräsvirta [1994] considère deux fonctions de transition : la fonction logistique et la fonction exponentielle. La figure 1 illustre le profil des deux fonctions de transition. Les modèles ainsi obtenus sont respectivement nommés LSTAR (*Logistic Smooth Transition AutoRegressive*) et ESTAR (*Exponential Smooth Transition AutoRegressive*).

Figure 1: Fonctions de transition



Le modèle LSTAR La fonction F est une fonction logistique donnée par :

$$F(Z_t; \gamma, s) = [1 + \exp(-\gamma(Z_t - s))]^{-1}$$

Ainsi, si $|Z_t - s|$ est grand et $Z_t < s$, la fonction de transition tend vers 0 et le processus décrit une période de récession. Y_t est engendré par le modèle linéaire autorégressif :

$$Y_t = \phi_0^{(1)} + \phi^{(1)'} X_t + \epsilon_t$$

Si $|Z_t - s|$ est grand mais $Z_t > s$, la fonction de transition tend vers 1 et le processus décrit une expansion. Y_t est engendré par :

$$Y_t = \phi_0^{(2)} + \phi^{(2)'} X_t + \epsilon_t$$

Des valeurs intermédiaires de la variable de transition Z_t conduisent à des combinaisons des modèles extrêmes. De plus, si le paramètre de lissage devient très grand, on obtient un modèle TAR puisque F devient une fonction indicatrice. Il s'agit donc d'une

généralisation du modèle TAR. Le modèle LSTAR décrit des situations où les phases de récession et d'expansion peuvent avoir des dynamiques différentes, mais la transition entre les deux régimes extrêmes est lisse. Ce modèle est ainsi capable d'engendrer des réalisations asymétriques.

Le modèle ESTAR La fonction F est une fonction exponentielle donnée par :

$$F(Z_t; \gamma, s) = \left[1 - \exp(-\gamma^*(Z_t - s^*)^2) \right]$$

Il s'agit d'une généralisation du modèle exponentiel autorégressif (EAR) de Ozaki [1978], où $\phi_0^{(1)} = \phi_0^{(2)} = 0$ et $s^* = 0$:

$$Y_t = (\phi_1^{(2)} + (\phi_1^{(1)} - \phi_1^{(2)})e^{-\gamma Y_{t-1}^2})Y_{t-1} + \dots + (\phi_p^{(2)} + (\phi_p^{(1)} - \phi_p^{(2)})e^{-\gamma Y_{t-p}^2})Y_{t-p} + u_t$$

Quand $|Z_t - s|$ est grand (expansion ou récession), la fonction F tend vers 1, et le modèle est décrit par les coefficients $\phi^{(2)}$. Alors que quand la variable de transition est proche du seuil s , la fonction F tend vers 0, et le modèle est plutôt décrit par les coefficients $\phi^{(1)}$. Le modèle ESTAR est une généralisation du modèle exponentiel introduit par Haggan et Ozaki [1981]. Il représente des économies où les récessions et les expansions ont plutôt des dynamiques similaires, s'opposant à la dynamique du régime intermédiaire. Un modèle ESTAR peut alors représenter une économie dont le passage d'un taux de croissance élevé à un taux de croissance normal se fait de la même manière que celui d'une faible croissance (voir négative) à une croissance moyenne. Il peut être vu comme une généralisation d'un modèle TAR à 3 régimes, où les deux régimes extrêmes seraient contraints par la même dynamique.

Un tel modèle permet de générer des oscillations auto-entretenues, indépendantes des conditions initiales, qui deviennent asymptotiquement périodiques, appelées cycles limites. Haggan et Ozaki [1981] donnent les conditions nécessaires à l'existence d'un cycle limite dans un modèle EAR, qui peuvent être étendues dans le cas du modèle ESTAR :

- toutes les racines de $z^k - \phi_1^{(2)}z^{k-1} - \phi_2^{(2)}z^{k-2} - \dots - \phi_k^{(2)} = 0$ sont à l'intérieur du disque unité.
- certaines racines de $z^k - \phi_1^{(1)}z^{k-1} - \phi_2^{(1)}z^{k-2} - \dots - \phi_k^{(1)} = 0$ sont à l'extérieur du disque unité.

Une condition suffisante à l'existence d'un cycle limite est :

$$\frac{1 - \sum_{i=1}^k \phi_i^{(2)}}{\sum_{i=1}^k (\phi_i^{(1)} - \phi_i^{(2)})} > 1 \quad \text{ou} \quad < 0$$

Cette dernière condition est requise pour supprimer toute éventualité d'un point singulier stable (voir Priestley [1989]).

3.1.2 Tests de linéarité

De même que dans le cas des modèles à seuil à transition brutale, les tests de linéarité contre les modèles STAR, ainsi que la spécification de ces modèles reposent sur un modèle linéaire de référence. Le test de linéarité consiste ensuite à tester ce modèle linéaire contre l'alternative fournie par le modèle STAR, le modèle linéaire étant emboîté dans le modèle non linéaire. Le modèle linéaire est en effet une restriction du modèle non linéaire, puisqu'il est obtenu quand les coefficients de la partie non linéaire sont nuls. Le statistique de test adéquate est alors un des tests asymptotiques suivants : test de Wald, test du rapport de vraisemblance ou test du multiplicateur de Lagrange (noté test LM par la suite). Parmi ces tests, le test LM est retenu puisqu'il ne nécessite que l'estimation du modèle sous l'hypothèse nulle, le modèle linéaire.

De même que dans la spécification des modèles TAR, Teräsvirta [1994] propose d'utiliser le test de linéarité pour déterminer la variable de transition, ou le paramètre de délai dans le cas univarié. Ainsi, les tests de linéarité sont construits pour toutes les variables de transition possibles, et on retient ensuite la variable de transition qui minimise le niveau de la statistique.

Tests de linéarité contre un modèle LSTAR

Afin d'alléger les notations et les calculs dans l'exposé des tests de linéarité, on peut réécrire le modèle LSTAR de manière à ne faire intervenir qu'une fois la fonction de transition :

$$Y_t = \pi_{10} + \pi'_1 X_t + (\pi_{20} + \pi'_2 X_t) \times \left\{ [1 + \exp(-\gamma(Z_t - s))]^{-1} - \frac{1}{2} \right\} + \epsilon_t \quad (5)$$

Ainsi $\pi_{10} = \phi_0^{(1)}$, $\pi'_1 = \phi^{(1) \prime}$, $\pi_{20} = \phi_0^{(2)} - \phi_0^{(1)}$, $\pi'_2 = \phi^{(2) \prime} - \phi^{(1) \prime}$.

On soustrait $\frac{1}{2}$ à la fonction logistique pour la construction du test de linéarité, ceci afin que la fonction de transition évaluée en $\gamma = 0$ soit nulle. Ainsi, sous l'hypothèse nulle $\gamma = 0$, on a le modèle linéaire avec les coefficients $\{\pi_{10}, \pi'_1\}$.

On remarque que l'hypothèse $\pi_{20} = \pi'_2 = 0$ conduit aussi à la linéarité, et peut donc être aussi retenue comme hypothèse nulle. Le fait d'avoir le choix pour la définition de l'hypothèse nulle de linéarité suggère une non-identifiabilité des paramètres. En effet si on retient comme hypothèse nulle : $H_0 : \gamma = 0$, les paramètres π_{20} , π'_2 , d et s ne sont pas identifiables sous H_0 , puisqu'ils peuvent prendre n'importe quelle valeur sans changer celle de la log-vraisemblance. Le problème de test en présence de paramètres de nuisance présents uniquement sous l'hypothèse alternative a été étudié par Davies [1977]. La solution n'est étudiée d'une manière générale que dans le cadre du test LM. L'idée est de construire la statistique de test LM en fonction des paramètres de nuisance, supposés fixes, et de maximiser ensuite la statistique par rapport à ces paramètres. On obtient ainsi une statistique conservatrice, dans le sens où on ne favorise pas le rejet de l'hypothèse nulle.

L'hypothèse de linéarité retenue par Teräsvirta [1994] est $H_0 : \gamma = 0$ contre l'alternative fournie par le modèle LSTAR (5) avec $\gamma > 0$. Si on note $\tau = (\tau_1', \tau_2)'$ où $\tau_1 = (\pi_{10}, \pi_1')$ et $\tau_2 = \gamma; (\hat{\tau}_1', 0)'$, l'estimation de ce vecteur de paramètres sous l'hypothèse nulle; $w_t = w_t(\tau) = ((\partial\epsilon_t/\partial\tau_1)', (\partial\epsilon_t/\partial\tau_2)')' = (w_{1t}', w_{2t}')'$, le vecteur de dérivées des résidus par rapport aux paramètres; $\hat{w}_t = (\hat{w}_{1t}', \hat{w}_{2t}')'$ ce vecteur sous l'hypothèse nulle, la statistique de test LM est donnée par¹⁴ :

$$LM = \hat{\sigma}^2 \left(\sum \hat{\epsilon}_t \hat{w}_{2t} \right)^2 \left[\sum \hat{w}_{2t}^2 - \sum \hat{w}_{2t} \hat{w}_{1t}' \left(\sum \hat{w}_{1t}' \hat{w}_{1t}' \right)^{-1} \sum \hat{w}_{1t}' \hat{w}_{2t} \right]$$

où $\hat{\sigma}^2 = (1/T) \sum \hat{\epsilon}_t^2$, $\hat{\epsilon}_t = Y_t - \pi_{10} - \pi_1' X_t$. Cette statistique suit une distribution du χ^2 à un degré de liberté.

Saikkonen et Luukkonen [1988] font remarquer que $LM = TR^2$ où R^2 est le coefficient de corrélation dans la régression de $\hat{\epsilon}_t$ sur \hat{w}_{1t}' et \hat{w}_{2t} . Ceci fournit une procédure de test simple à mettre en application.

Afin d'obtenir la statistique LM du test $\gamma = 0$ contre un modèle LSTAR, il faut calculer \hat{w}_{1t}' et \hat{w}_{2t} , les dérivées des résidus ϵ_t par rapport aux paramètres π_{10} , π_1' et γ évaluées sous H_0 , les paramètres de nuisance étant fixés :

$$\begin{aligned} \left. \frac{\partial\epsilon_t}{\partial\pi_{10}} \right|_{\gamma=0} &= -1, \\ \left. \frac{\partial\epsilon_t}{\partial\pi_1'} \right|_{\gamma=0} &= -X_t, \\ \left. \frac{\partial\epsilon_t}{\partial\gamma} \right|_{\gamma=0} &= -(1/4)[\pi_{20}(Z_t - s) - s\pi_2' X_t + \pi_2' X_t Z_t] \end{aligned}$$

Ainsi $\hat{w}_{1t} = -(1, X_t)'$ et \hat{w}_{2t} dépend du vecteur de paramètres de nuisance $\pi = (\pi_{20}, \pi_2', d, s) : \hat{w}_{2t}(\pi) = -(1/4)[\pi_{20}(Z_t - s) - s\pi_2' X_t + \pi_2' X_t Z_t]$.

En pratique, le test peut alors se construire de la manière suivante :

- i) Régresser Y_t sur $(1, X_t)$, en déduire les résidus estimés $\hat{\epsilon}_t$ et $SCR_0 = \sum \hat{\epsilon}_t^2$.
- ii) Effectuer la régression :

$$\hat{\epsilon}_t = \hat{w}_{1t}' \tilde{\beta}_1 + \hat{w}_{2t}(\pi) \tilde{\beta}_2 + u_t \quad (6)$$

Construire la somme des carrés des erreurs $SCR(\pi)$.

- iii) Construire la statistique $LM(\pi) = T \times \frac{SCR_0 - SCR(\pi)}{SCR(\pi)}$, dont la distribution asymptotique est un $\chi^2(1)$ sous $H_0 : \tilde{\beta}_2 = 0$ ($\gamma = 0$) mais dépend du vecteur de paramètres π . A l'instar de Davies [1977], on retient la statistique conservatrice suivante :

$$LM_1 = \sup_{\pi} LM(\pi) = T \times \frac{SCR_0 - \inf_{\pi} SCR(\pi)}{SCR_0}$$

¹⁴Voir Saikkonen et Luukkonen [1988] et Godfrey [1988] pour une description des tests du multiplicateur de Lagrange.

Cette statistique est difficile à calculer puisqu'il faut calculer $SCR(\pi)$ pour toutes les valeurs possibles de tous les paramètres de nuisance. De plus, la distribution est inconnue sous l'hypothèse nulle. Néanmoins Teräsvirta [1994] montre qu'en organisant les régresseurs de l'équation (6), on peut la réécrire comme :

$$\hat{\epsilon}_t = -\delta_{10} - \delta_{11}X_t + \delta_2X_tZ_t + v_t \quad (7)$$

avec : $\delta_{10} = \pi_{10} - (s/4)\pi_{20}$, $\delta_{11} = \pi_1 - (s/4)\pi_2 + (1/4)\pi_{20}e_d$, $\delta_2 = \pi_2/4$ où e_d est le vecteur pour lequel la d ième composante est 1 et le reste est constitué de zéro (ceci dans le cas univarié). Ainsi, minimiser la somme des carrés des erreurs de la régression (7) par rapport à δ_{10} , δ_{11} et δ_2 donne $\inf_{\pi} SCR$. De plus, l'hypothèse nulle étant $\delta_2 = 0$ (m contraintes), la statistique LM_1 est asymptotiquement distribuée selon un $\chi^2(m)$ sous H_0 .

Il convient de remarquer qu'alors que la linéarité implique $\delta_2 = 0$ dans (7), δ_2 est indépendant de π_{20} . Ainsi, dans une situation où la non-linéarité serait essentiellement due à π_{20} , un test basé sur (7) aurait peu de puissance. La solution apportée par Luukkonen et al. [1988a] consiste à remplacer $F(z)$, où $z = \gamma(Z_t - s)$, par son développement de Taylor à l'ordre 3¹⁵ :

$$T_3(z) = g_1z + g_3z^3, \text{ où } g_1 = \left. \frac{\partial F}{\partial z} \right|_{z=0}, g_3 = \left. \frac{1}{6} \frac{\partial^3 F}{\partial z^3} \right|_{z=0}$$

Le modèle devient, en réarrangeant les termes :

$$Y_t = \beta_{10} + \beta'_1X_t + \beta'_2X_tZ_t + \beta'_3X_tZ_t^2 + \beta'_4X_tZ_t^3 + v_t \quad (8)$$

avec

$$\begin{aligned} \beta_{10} &= \pi_{10} - \pi_{20}g_1\gamma s - \pi_{20}g_3\gamma^3s^3 \\ \beta_1 &= \pi_1 - \pi_2g_1\gamma s - \pi_2g_3\gamma^3s^3 + \pi_{20}g_1\gamma e_d + 3\pi_{20}g_3\gamma^3s^2 \\ \beta_2 &= \pi_2g_1\gamma + 3\pi_2g_3\gamma^3s^2 - 3\pi_{20}g_3\gamma^3se_d \\ \beta_3 &= -3\pi_2g_3\gamma^3s + \pi_{20}g_3\gamma^3e_d \\ \beta_4 &= g_3\gamma^3\pi_2 \end{aligned}$$

Le test de linéarité $\gamma = 0$ revient alors à tester :

$$H'_0 : \beta_{2j} = \beta_{3j} = \beta_{4j} = 0 \quad (j = 1, \dots, m) \quad (9)$$

dans la régression (8).

On peut noter que ce test est équivalent au test décrit précédemment dans la régression (7), mais dans une régression augmentée des termes $X_tZ_t^2$ et $X_tZ_t^3$. Ainsi, la statistique de test se construit de la manière suivante :

¹⁵Le développement de Taylor à l'ordre 1 ne permet pas de résoudre ce problème, et la dérivée seconde de $F(z)$ par rapport à z évaluée en $z = 0$ est nulle. Il faut donc avoir recours au développement à l'ordre 3. Le développement de Taylor n'est qu'une approximation de la fonction de transition. Cependant, il permet de mettre en évidence la présence de non-linéarités si les coefficients des termes croisés sont significativement non nuls. Dans ce cas, le modèle non linéaire sera estimé globalement.

- i) Régresser Y_t sur $(1, X_t)$, en déduire les résidus estimés $\hat{\epsilon}_t$ et $SCR_0 = \sum \hat{\epsilon}_t^2$.
- ii) Effectuer la régression (8). En déduire la somme des carrés des résidus SCR_2 .
- iii) Construire la statistique $LM_2 = T \times \frac{SCR_0 - SCR_2}{SCR_0}$. La distribution asymptotique de LM_2 est un $\chi^2(3m)$.

Alors que ce test a été construit afin d'être puissant contre des alternatives où la non-linéarité proviendrait de π_{20} , ce paramètre intervient dans β_3 . Ainsi, Teräsvirta [1994] propose un test plus parsimonieux, qui consiste à tester de la même manière la linéarité dans la régression (7) augmentée seulement du régresseur Z_t^3 .

Tests de linéarité contre un modèle ESTAR

De même que pour le modèle LSTAR, on peut réécrire le modèle ESTAR de manière à ne faire intervenir qu'une fois la fonction de transition :

$$Y_t = \theta_{10} + \theta'_1 X_t + (\theta_{20} + \theta'_2 X_t) \times \left[\left(1 - \exp(-\gamma^*(Z_t - s^*)^2) \right) \right] + v_t \quad (10)$$

avec $\theta_{10} = \phi_0^{(1)}$, $\theta'_1 = \phi^{(1) '}$, $\theta_{20} = \phi_0^{(2)} - \phi_0^{(1)}$, $\theta'_2 = \phi^{(2) '} - \phi^{(1) '}$.

Teräsvirta [1994] retient $\gamma^* = 0$ comme hypothèse nulle de linéarité dans le modèle (10). En effet, si $\gamma^* = 0$, la fonction de transition est nulle, et le modèle est linéaire avec les paramètres $\{\theta_{10}, \theta'_1\}$.

De la même manière que dans le modèle LSTAR, on construit la statistique de test de $H_0 : \gamma^* = 0$ à partir du calcul des dérivées des résidus v_t par rapport aux paramètres $\tau = (\tau'_1, \tau_2)'$ où $\tau_1 = (\theta_{10}, \theta'_1)'$ et $\tau_2 = \gamma^*$, en supposant que les paramètres de nuisance $\pi = \theta_{20}, \theta'_2, d, s^*$ sont donnés.

On obtient :

$$\begin{aligned} \left. \frac{\partial v_t}{\partial \theta_{10}} \right|_{\gamma^*=0} &= -1, \\ \left. \frac{\partial v_t}{\partial \theta'_1} \right|_{\gamma^*=0} &= -X_t, \\ \left. \frac{\partial v_t}{\partial \gamma^*} \right|_{\gamma^*=0} &= -(Z_t - s^*)^2 (\theta_{20} + \theta'_2 X_t) \\ &= -(\theta'_2 X_t Z_t^2 + \theta_{20} Z_t^2 - 2s^* \theta'_2 X_t Z_t + s^{*2} \theta'_2 X_t - 2s^* \theta_{20} Z_t + s^{*2} \theta_{20}) \end{aligned}$$

Ainsi $\hat{w}_{1t} = -(1, X_t)'$ et \hat{w}_{2t} dépend du vecteur de paramètres de nuisance π .

Ceci suggère la régression auxiliaire suivante :

$$\hat{v}_t = \tilde{\beta}'_1 \hat{w}_{1t} + \gamma'_2 X_t Z_t + \gamma'_3 X_t Z_t^2 + e'_t \quad (11)$$

où e'_t est un terme d'erreur et $\tilde{\beta}_1 = (\beta_{10}, \beta'_1)'$ avec

$$\begin{aligned} \beta_{10} &= \theta_{10} - s^{*2} \theta_{20} \\ \beta'_1 &= \theta'_1 - s^{*2} \theta'_2 + 2s^* \theta_{20} e_d \\ \gamma_2 &= 2s^* \theta'_2 - \theta_{20} e_d \\ \gamma_3 &= -\theta_2 \end{aligned}$$

Ainsi l'hypothèse nulle de linéarité est $H_0 : \gamma_2 = \gamma_3 = 0$ et sous cette hypothèse, la statistique de test est $LM_3 = T \times \frac{SCR_0 - SCR_3}{SCR_0}$ où SCR_3 est la somme des carrés des résidus du modèle (11). LM_3 a une distribution asymptotique du $\chi^2(2m)$.

Ce test est le premier des trois tests décrits par Luukkonen et al. [1988a] quand le paramètre de délai d est inconnu. Ils le construisent à partir d'un développement de Taylor à l'ordre un. Cependant, comme on l'a noté précédemment, dans certaines situations, cette approximation est inadéquate (pas de puissance contre certaines alternatives) et une approximation à l'ordre 3 est nécessaire pour construire le test de linéarité (voir Luukkonen et al. [1988a]).

Choix entre le modèle LSTAR et le modèle ESTAR

Une dernière étape dans la spécification du modèle non linéaire est de choisir entre les modèles LSTAR et ESTAR. Il suffit de comparer les deux régressions auxiliaires associées respectivement aux tests de linéarité contre les modèles LSTAR et ESTAR (voir Teräsvirta [1994]). Ces deux régressions, données par les équations (8) et (11), diffèrent puisque le terme en $X_t Z_t^3$ n'est présent que dans la régression auxiliaire associée au modèle LSTAR. Teräsvirta [1994] propose alors de construire la séquence de tests suivante dans le modèle le plus général, équation (8) :

$$\text{Test } F_4 : \begin{cases} H_{01} : \beta_4 = 0 & \text{dans (8)} \\ H_{11} : \beta_4 \neq 0 & \text{dans (8)} \end{cases} \quad (12)$$

$$\text{Test } F_3 : \begin{cases} H_{02} : \beta_3 = 0 / \beta_4 = 0 & \text{dans (8)} \\ H_{12} : \beta_3 \neq 0 / \beta_4 = 0 & \text{dans (8)} \end{cases} \quad (13)$$

$$\text{Test } F_2 : \begin{cases} H_{03} : \beta_2 = 0 / \beta_3 = \beta_4 = 0 & \text{dans (8)} \\ H_{13} : \beta_2 \neq 0 / \beta_3 = \beta_4 = 0 & \text{dans (8)} \end{cases} \quad (14)$$

L'intuition de cette méthode repose sur la comparaison des coefficients des régressions (8) et (11).

- Si le vrai modèle est ESTAR, alors il n'y a pas de termes au cube. Ainsi rejeter H_{01} peut être interprétée comme un rejet du modèle ESTAR, en faveur du modèle LSTAR.
- On a $\gamma_3 = 0$ si et seulement si $\theta_2 = 0$, ce qui est un cas particulier et peu fréquent du modèle ESTAR, alors que si le vrai modèle est LSTAR, on peut avoir $\beta_3 = 0$ si $s = \pi_{20} = 0$. Ainsi si l'hypothèse H_{02} n'est pas rejetée, on peut espérer que le vrai modèle soit LSTAR. En revanche, un rejet de cette hypothèse n'informe pas sur le modèle à retenir.
- Il semble que si le vrai modèle est LSTAR, β_2 sera différent de zéro et H_{03} sera rejetée. Ainsi rejeter H_{03} après avoir accepté H_{02} conclurait en faveur du modèle LSTAR. De plus, il peut être possible d'accepter H_{03} si le modèle est ESTAR et $s^* = \theta_{20} = 0$.

La méthode proposée par Teräsvirta [1994] est de tester la linéarité dans le modèle le plus général, c'est-à-dire tester l'hypothèse nulle (9) dans la régression (8). Si la linéarité est rejetée, on choisit entre les alternatives LSTAR et ESTAR en construisant les trois tests comme des tests de Fisher et on compare les niveaux de rejet. En effet, si le modèle est LSTAR, H_{01} et H_{03} sont rejetées plus fortement que H_{02} . Si le modèle ESTAR, la situation opposée est attendue.

Teräsvirta [1994] analyse les performances de cette procédure par simulations. Il conclut que cette méthode de sélection donne de bons résultats quand le vrai modèle est LSTAR. Les résultats sont moins bons quand le vrai modèle est ESTAR, et que les observations ne sont pas distribuées de manière symétrique autour du seuil. Cependant, il est important de noter que dans ce cas, les deux modèles, LSTAR et ESTAR sont des substituts très proches. Dans ce cas, il est important d'estimer les deux modèles et de comparer les résultats d'estimation.

3.1.3 L'estimation des modèles STAR

L'estimation des paramètres des modèles STAR est obtenue par la méthode des moindres carrés non linéaires. Cette méthode est équivalente à la maximisation de la fonction de vraisemblance si les erreurs sont normalement distribuées. Si le modèle s'écrit :

$$Y_t = g(\alpha, X_t) + u_t$$

le problème est alors de minimiser

$$Q_T(\alpha) = \sum_{t=1}^T \{Y_t - g(\alpha, X_t)\}^2$$

par rapport à α .

Tong [1990] donne les conditions de convergence et de normalité asymptotique des estimateurs, mais elles sont difficiles à vérifier en pratique, de telle sorte que les écart-types estimés et les intervalles de confiance qui en résultent sont à utiliser avec prudence.

Dans le modèle LSTAR, Teräsvirta [1994] explique que l'estimation jointe du vecteur de paramètres est difficile. L'estimation de γ est la plus problématique, car une valeur élevée de ce paramètre empêche l'algorithme de converger (surtout si s est en même temps très faible) : ceci est imputable à l'utilisation d'une procédure numérique¹⁶. Teräsvirta [1994] propose de normaliser le paramètre de lissage en le divisant par l'écart-type de la variable de transition. Ceci permet en effet de prendre comme condition initiale des valeurs proches de 1, alors que sans cette normalisation, le paramètre de lissage peut varier énormément, ce qui rend difficile le choix de la condition initiale.

Pour le modèle ESTAR, la difficulté d'estimer simultanément l'ensemble des paramètres est due à une forte corrélation négative entre l'estimateur de γ^* et celui de $\phi^{(2)}$,

¹⁶On utilise la méthode du simplexe en première étape d'estimation. Elle permet en effet une convergence plus lente que les méthodes du gradient. Les résultats définitifs sont finalement obtenus après application de méthodes du gradient.

empêchant alors la convergence de l'algorithme. Une solution est alors dans ce cas, de diviser le paramètre de lissage par la variance de la série. Cette normalisation permet de trouver plus facilement une valeur initiale pour le paramètre de lissage (1 sera souvent une valeur initiale satisfaisante). Afin de trouver des valeurs initiales pour les autres paramètres, le modèle linéaire autorégressif est estimé. Ces paramètres autorégressifs sont retenus comme conditions initiales pour l'estimation du modèle STAR contraint par $\gamma = 1$. De nouveau, ces estimations fournissent les conditions initiales pour l'estimation globale du modèle non linéaire¹⁷.

La convergence du paramètre de lissage peut être lente et l'estimation de l'écart-type de ce paramètre sera souvent élevée. Cela indique que la valeur de ce paramètre est grande et que la transition entre les régimes est assez rapide (modèle se rapprochant des modèles TAR).

Validation Le modèle non linéaire étant estimé, il importe d'en valider les résultats. Teräsvirta [1994] propose de vérifier tout d'abord si le seuil s appartient bien à l'échelle de la série et si les écart-type des paramètres estimés (sauf celui de γ) ne sont pas trop grands (sinon il y a risque de redondances dans les paramètres). On peut valider l'estimation du modèle par des tests de spécification sur les résidus estimés du modèle : test de normalité des résidus de Jarque et Bera, le test d'absence d'hétéroscédasticité conditionnelle d'ordre k (ARCH) de McLeod et Li [1983] (test LM), un test d'indépendance des erreurs¹⁸. Eitrheim et Teräsvirta [1996] proposent aussi un test de mauvaise spécification de la non-linéarité, consistant à tester l'absence de non-linéarités dans les résidus estimés. Enfin, ils proposent un test de constance des paramètres du modèle STAR.

4 Modèles à changement de régime markovien

Les modèles à changement de régimes, introduits en économie par Hamilton [1989], sont traditionnellement utilisés pour prendre en compte des ruptures, comme une guerre ou des crises, ou de nouvelles politiques économiques, qui peuvent modifier fortement l'évolution des variables. En effet, les paramètres du modèle peuvent ne changer qu'occasionnellement quand l'état du système change.

4.1 Présentation du modèle

Hamilton [1989] propose de rendre compte de la différence de dynamique durant les périodes d'expansion et les périodes de récession, en supposant que le changement

¹⁷Voir Schlittgen [1997] pour une discussion sur l'estimation des modèles STAR. Cet auteur préconise l'utilisation d'une grille quand l'estimation est difficile. Voir aussi Teräsvirta [1994] et Granger et Teräsvirta [1993].

¹⁸Le test du portmanteau de Ljung et Box [1978] n'est pas applicable car sa distribution asymptotique n'est pas connue quand on l'applique sur les résidus du modèle STAR.

de régime est gouverné par un processus de Markov à plusieurs états ($s_t = 0, 1$). Il caractérise le taux de croissance du PIB aux Etats-Unis y_t par un modèle à deux régimes :

$$\begin{aligned} y_t &= \alpha_0 + \alpha_1 s_t + z_t \\ z_t &= \phi_1 z_{t-1} + \phi_2 z_{t-2} + \phi_3 z_{t-3} + \phi_4 z_{t-4} + \varepsilon_t \end{aligned}$$

Son intérêt se porte plus particulièrement sur le changement dans le taux de croissance moyen selon la phase du cycle depuis la fin de la seconde Guerre Mondiale. Il estime que le taux de croissance est égal à $\alpha_0 = -0.4\%$ par trimestre quand $s_t = 0$ et $\alpha_0 + \alpha_1 = 1.2\%$ quand $s_t = 1$.

L'approche retenue ici est différente¹⁹ puisque l'on ne recherche pas des ruptures dans la croissance, mais que l'on veut identifier des régimes aux dynamiques distinctes, afin de rendre compte d'éventuelles asymétries. A la différence des modèles à seuil, on n'a pas à choisir la variable qui provoque le changement de régime parmi les variables du modèle. En effet, on suppose, dans le modèle Markovien, que le changement de régime est gouverné par une variable exogène, inobservable, pour laquelle on spécifie une loi de probabilité. L'estimation des paramètres permet d'identifier les régimes *a posteriori*. L'avantage de la procédure de Hamilton est qu'elle permet de plus d'identifier les points de retournements du cycle, et de retrouver une datation très proche de celle obtenue par le *National Bureau of Economic Research* (NBER) :

“Il est intéressant que la procédure a-théorique des modèles à changement de régimes Markoviens obtienne les mêmes dates que la procédure du NBER. Ceci permet de renforcer le fait que les récessions correspondent à des épisodes caractérisés par des dynamiques distinctes des périodes d'expansion”. (Hamilton [1993], page 257).

Un autre avantage de cette méthode est l'obtention des probabilités de transition entre les régimes, qui permet le calcul des durées espérées de chaque régime, ce qui fournit un indicateur de la persistance des régimes.

Afin d'autoriser la présence d'asymétries dans la dynamique des variables, le modèle initialement proposé par Hamilton [1989] est étendu. Le modèle à changement de régimes Markoviens adopté ici définit alors différents régimes autorégressifs²⁰ qui dépendent de l'état du système. La modélisation de la variable endogène y_t est donc la suivante :

$$y_t = \mu^{s_t} + \phi^{s_t} Y_{t-1} + \sigma^{s_t} \varepsilon_t \quad (15)$$

où ε_t est $\mathcal{N}(0, 1)$, $\phi^{s_t} = (\phi_1, \dots, \phi_p)^{s_t}$ et $Y_{t-1} = (y_{t-1}, \dots, y_{t-p})'$.

¹⁹Nous considérons en effet des variables stationnaires.

²⁰Voir aussi Warne [1995], Kugler [1996] sur la généralisation du modèle de Hamilton.

Le modèle est linéaire par régime²¹. Tous les coefficients, relatifs à la moyenne et à la partie autorégressive, peuvent alors différer entre les régimes, ce qui n'est pas le cas dans le modèle initialement proposé par Hamilton. Ceci permet non seulement de rendre compte de différences dans le niveau de l'endogène selon le régime, mais aussi de différences dans la dynamique interne au régime (par exemple différence de persistance, ou différence dans l'amplification des chocs)²².

La variable s_t est supposée être une variable latente inobservable, qui représente l'état de l'économie. Elle prend les valeurs $1, 2, \dots, S$, où S représente le nombre d'états. Par exemple, si $S = 2$, la variable y_t est caractérisée par deux régimes : un premier où la dynamique autorégressive est donnée par $\mu^{(1)} + \phi^{(1)}Y_{t-1} + \sigma^{(1)}\varepsilon_t$ et un second où la dynamique est $\mu^{(2)} + \phi^{(2)}Y_{t-1} + \sigma^{(2)}\varepsilon_t$. Ici, la dynamique de l'endogène est supposée ne dépendre que de l'état courant du système, contrairement à la spécification retenue par Hamilton [1989].

De plus, la variable s_t est supposée suivre un processus de Markov du premier ordre. Elle est ainsi caractérisée par des probabilités de transition entre les différents états, qui ne dépendent que de l'état à la date précédente. Ces probabilités de transition, notées p_{ij} sont comprises entre 0 et 1 et sont telles que : $P(s_t = j | s_{t-1} = i, s_{t-2} = k, \dots) = P(s_t = j | s_{t-1} = i) = p_{ij}$ avec $\sum_{j=1}^S p_{ij} = 1 \forall i$. Ainsi, si l'économie était dans l'état 1 à la période précédente, la probabilité de passer dans l'état 2 est donnée par p_{12} ²³.

Cette représentation est très proche de celle fournie par les modèles à seuil à transition brutale (modèles TAR), étudiés dans la première section. En effet, les deux familles de modèles, modèles TAR et modèles Markoviens, supposent que la dynamique de l'endogène n'est pas la même selon le régime où l'économie se situe. Cependant, ils diffèrent dans le choix de la variable qui régit le changement de régime. Alors que la variable de transition est une variable observable du modèle dans le cas des modèles TAR, c'est une variable inobservable qui remplit ce rôle dans les modèles Markoviens. Le choix de la variable de transition dans les modèles TAR est laissé au soin de l'économètre ou des tests de linéarité. Dans les modèles Markoviens, le changement de régime est gouverné par une variable extérieure, ou exogène, non observable. Une fois la loi de cette variable spécifiée, elle est ensuite définie telle que la dynamique de la variable étudiée soit différente selon l'état du système. Ainsi, l'estimation de l'ensemble des paramètres

²¹Ce modèle est globalement non linéaire, à cause de la présence de la variable s_t . Voir Hamilton [1994a], chapitre 22, pour plus de détails sur la nature des non-linéarités engendrées par ce modèle.

²²Voir l'application de Kugler [1996] sur les taux d'intérêt. Voir aussi Hénin et P. [1998] sur les interactions entre les régimes budgétaires et les phases conjoncturelles.

²³Les probabilités de transition sont ici supposées constantes. Filardo [1994] généralise le modèle proposé par Hamilton [1989] en introduisant des probabilités de transition qui varient au cours du temps, selon la phase du cycle. Elles dépendent en fait de variables explicatives, ceci afin d'augmenter la corrélation entre la variable étudiée et le cycle. L'amélioration des résultats relativement à ceux du modèle standard de Hamilton [1989] n'est pas très élevée. De plus, cette spécification impose plus de paramètres à estimer. Voir aussi Garcia et Schaller [1995] pour une application sur les effets des politiques monétaires.

du modèle - paramètres autorégressifs de chaque régime et paramètres associés à la loi de la variable inobservable - permet de mener une inférence sur les probabilités d'être plutôt dans un état à une date donnée. Le modèle peut être utilisé en prévision même si l'état n'est pas directement observable, puisque l'on connaît la loi de probabilité de cette variable²⁴. L'avantage des modèles Markoviens est de ne pas avoir à choisir la variable de transition parmi l'ensemble des variables du modèle. Les résultats d'estimation de ces deux familles de modèles (TAR et Markoviens) seront comparés.

Le modèle à changement de régimes Markoviens est pertinent pour étudier les différents types d'asymétries conjoncturelles. En effet, supposons qu'il n'y ait que deux états : $s_t = 1$ ou $s_t = 2$. Le régime 1 correspond par exemple à des périodes dans lesquelles la variable endogène converge vers un niveau élevé. Ce régime est alors caractérisé par un taux de croissance de l'endogène positif. Le régime 2 est tel que la dynamique converge vers un niveau bas, et le taux de croissance est alors négatif. Ainsi, ces régimes permettent de rendre compte d'asymétries entre les phases de croissance et de décroissance si les propriétés de la dynamique du premier régime ne sont pas les mêmes que celles du second régime. La comparaison des coefficients autorégressifs dans chaque régime conduit à détecter d'éventuelles différences dans les ajustements à la hausse et à la baisse. Comme le souligne Hamilton²⁵, ces régimes ainsi identifiés permettent de caractériser les points de retournements. L'asymétrie dans l'amplitude des pics et des creux peut alors être caractérisée. L'estimation des probabilités de transition permet finalement de calculer la durée moyenne des différents régimes, et de mettre ainsi en évidence l'asymétrie dans la durée des régimes. Sachant que l'économie était dans le régime i , la durée espérée d'un régime est donnée par : $\sum_{k=1}^{\infty} k p_{ii}^{k-1} (1-p_{ii}) \rightarrow (1-p_{ii})^{-1}$.

Il est intéressant de noter que le modèle à changement de régimes Markoviens, donné par l'équation (15), comprend comme cas particulier le modèle linéaire autorégressif, quand $p_{ii} = 0$. De plus, cette spécification permet aussi de rendre compte de changements permanents, si $p_{12} = 0$ par exemple. Si tel n'est pas le cas, la probabilité de rester dans un régime est strictement inférieure à 1 et les probabilités de transition entre les régimes sont strictement supérieures à 0. Il est essentiel de remarquer que ce type de modèle peut rendre compte de la dynamique d'une variable même s'il n'y a qu'un seul changement observé. En effet, la précision de l'inférence concernant les paramètres d'un régime dépend du nombre d'observations générées par ce régime, et non pas du nombre de changements de régimes. L'idée est que chaque nouvelle observation générée par un régime fournit une nouvelle information sur la probabilité de rester

²⁴Dans les modèles Markoviens, les changements de régimes ne sont pas déterministes, comme c'est le cas des modèles de changement structurels à la Perron [1989]. Dans ces derniers, la rupture n'est pas supposée pouvoir se reproduire, et les prévisions ne prennent pas en compte d'autres changements.

²⁵Voir Hamilton [1989], [1993], [1994b]. Consulter de plus Goodwin [1993], qui estime un modèle à changement de régimes Markoviens, similaire à celui proposé par Hamilton [1989], sur différents pays occidentaux. Alors que les performances en prévision et les tests de spécification ne concluent qu'à une amélioration marginale relativement aux modèles linéaires, il montre que ce type de modélisation permet de détecter les points de retournements dans le cycle.

dans ce régime. Les probabilités de transition sont ensuite déduites de l'estimation des probabilités de rester dans un régime.

En spécifiant la loi de probabilité de la variable d'état plutôt que d'imposer des dates particulières pour les changements de régimes, nous laissons les données nous révéler la nature et les conséquences des changements de régimes. Ainsi, le vecteur de paramètres $\theta = \{\mu^{(i)}, \phi^{(i)}, \sigma^{(i)}, p_{ij}\}$, ($i = 1, \dots, S$; $j = 1, \dots, S - 1$), caractérise la densité de probabilité des données. L'objectif est alors de trouver la valeur du vecteur de paramètres θ qui décrit au mieux les données. Cette estimation permettra alors de mener une inférence sur les observations associées à chaque état.

4.2 La procédure d'estimation

L'estimation des paramètres est obtenue par la méthode du maximum de vraisemblance. On cherche à maximiser la fonction de log-vraisemblance suivante :

$$\mathcal{L}(\theta) = \sum_{t=1}^T \log f(y_t | \mathcal{Y}_t; \theta)$$

où \mathcal{Y}_t est l'ensemble d'information disponible à la date t . $f(y_t | \mathcal{Y}_t; \theta)$ est la distribution – non conditionnelle à l'état – des observations. Puisque dans chaque régime, les observations sont normalement distribuées, avec des paramètres propres au régime, $f(y_t | \mathcal{Y}_t; \theta)$ est une combinaison linéaire de distributions normales, chacune étant associée à un état particulier et pondérée par la probabilité d'être dans l'état correspondant. A chaque date t , on évalue :

$$\begin{aligned} f(y_t | \mathcal{Y}_t; \theta) &= \sum_{j=1}^S P(y_t, s_t = j | \mathcal{Y}_t; \theta) \\ &= \sum_{j=1}^S P(s_t = j | \mathcal{Y}_{t-1}; \theta) f(y_t | s_t = j, \mathcal{Y}_{t-1}; \theta) \end{aligned}$$

où

- $f(y_t | s_t = j, \mathcal{Y}_{t-1}; \theta)$ est la densité de y_t conditionnellement à l'état j , c'est-à-dire une distribution normale $\mathcal{N}(\mu^{(j)} + \phi^{(j)} Y_{t-1}, \sigma^{(j)})$,
- $P(s_t = j | \mathcal{Y}_{t-1}; \theta) = \sum_{i=1}^S p_{ij} P(s_{t-1} = i | \mathcal{Y}_{t-1}; \theta)$ est la probabilité d'être dans l'état j à la date t conditionnellement à l'information en $t - 1$.

Ainsi, la fonction de vraisemblance est calculée à partir d'un algorithme itératif²⁶. Chaque itération de l'algorithme correspond aux étapes suivantes :

- (i) on entre $P(s_{t-1} = i | \mathcal{Y}_{t-1}; \theta)$,
- (ii) on calcule la valeur de la densité $f(y_t | \mathcal{Y}_t; \theta)$,

²⁶Il est couramment appelé filtre de Hamilton et est décrit plus en détails par Hamilton [1989].

(iii) on récupère $P(s_t = i | \mathcal{Y}_t; \theta) = \frac{P(y_t, s_t = i | \mathcal{Y}_t; \theta)}{f(y_t | \mathcal{Y}_t; \theta)}$, issue de la formule de Bayes. Cette probabilité est l'entrée de l'itération suivante.

Puisque la dynamique de y_t est supposée ne dépendre que de l'état courant s_t , et non pas de l'état dans lequel la variable endogène était à la période précédente (la densité ne dépend pas de $s_{t-1}, s_{t-2}, \dots, s_{t-p}$ comme dans le modèle standard de Hamilton [1989]), la densité ne tient pas compte des p états passés. Ainsi, elle n'implique pas le calcul d'intégrales sur l'ensemble des p états passés comme dans Hamilton [1989], ou de leur approximation comme dans Kim [1994].

La vraisemblance dépend d'une condition initiale. En effet, lors de la première itération, on doit donner comme entrée à l'algorithme la probabilité d'être dans un état donné à la période initiale. Elle peut être fixée *a priori* ou estimée. Dans le premier cas, on peut estimer tous les modèles Markoviens avec différentes conditions initiales, et retenir le modèle qui correspond à la plus grande vraisemblance (Phillips [1991]). Ceci requiert alors d'estimer plusieurs modèles Markoviens. Le second cas implique l'estimation de paramètres supplémentaires. Hamilton [1993] suggère plutôt de remplacer cette inconnue par la probabilité inconditionnelle d'être dans un état donné, *i.e.* $P(s_1 = i) = (1 - p_{jj}) / (2 - p_{11} - p_{22})$ avec $i \neq j$ dans le cas de $S = 2$. Ainsi, la probabilité d'être dans l'état i à la période initiale dépend des probabilités de transition. L'estimation des probabilités de transition tient alors compte de la condition initiale, et les estimations ne correspondent plus à des fréquences empiriques.

La maximisation de la vraisemblance est réalisée sous les contraintes $0 \leq p_{ij} \leq 1$ ²⁷ et $\sum_{j=1}^S p_{ij} = 1$. D'un point de vue pratique, la maximisation de la vraisemblance est réalisée comme suit :

- (i) la procédure de minimisation du simplexe, développé par Nelder et Mead²⁸ est mise en œuvre, elle implique en effet une progression lente dans la modification des paramètres, ce qui assure que les paramètres ne prennent pas de valeurs aberrantes,
- (ii) une méthode de gradient est ensuite appliquée en prenant comme conditions initiales pour les paramètres, les résultats issus de la première étape²⁹.

Les étapes (i) et (ii) sont répétées un grand nombre de fois. De plus, il est important de relancer la procédure d'estimation avec différentes conditions initiales, afin d'éviter les problèmes d'optima locaux.

²⁷Cette restriction est directement incorporée dans la fonction objectif par la transformation $x^2 / (1 + x^2)$.

²⁸Voir Press, Teukolsky, Vetterling et Flannery [1992] pour plus de détails sur cette méthode.

²⁹Une méthode alternative à l'optimisation numérique est donnée par l'algorithme EM (voir Hamilton [1990]). Les résultats présentés dans la suite ont été comparés avec ceux issus de l'application de l'algorithme EM. Les codes de programmation de l'algorithme EM ont été fournis par Warne [1995].

4.3 Inférence sur les états

Etant donné l'ensemble de paramètres θ qui décrivent au mieux les données, l'étape suivante consiste à identifier les régimes, c'est-à-dire à inférer l'état dans lequel le système se trouve à la date t . Puisque s_t n'est pas directement observable, l'inférence prend la forme d'une probabilité :

$$P(s_t = j | \mathcal{Y}_t; \theta) = \frac{P(y_t, s_t = j | \mathcal{Y}_t; \theta)}{f(y_t | \mathcal{Y}_t; \theta)}$$

Ces probabilités sont appelées “probabilités filtrées”. Elles sont basées sur l'information disponible jusqu'à la date t et sont calculées par l'algorithme décrit précédemment pour l'estimation de θ ³⁰.

D'autres probabilités peuvent être calculées. Il s'agit des “probabilités lissées”. Elles correspondent à l'inférence de l'état à la date t basée sur l'information de l'ensemble de l'échantillon, c'est-à-dire $P(s_t = j | \mathcal{Y}_T; \theta)$. Ainsi, étant donné $P(s_{t+k}, s_t | \mathcal{Y}_{t+k}; \theta)$ (obtenu pour $k = 1$ lors du calcul des probabilités filtrées, puis récursivement sur $k = 1, \dots, T - 1$), on obtient de manière récursive sur t :

$$P(s_{t+k+1}, s_t | \mathcal{Y}_{t+k+1}; \theta) = \frac{\sum_{s_{t+k}} P(s_{t+k}, s_t | \mathcal{Y}_{t+k}; \theta) P(s_{t+k+1} | s_{t+k}) f(y_{t+k+1} | s_{t+k+1}, s_{t+k}, \mathcal{Y}_{t+k}; \theta)}{f(y_{t+k+1} | \mathcal{Y}_{t+k}; \theta)}$$

Quand on atteint $P(s_T, s_t | \mathcal{Y}_T; \theta)$, on peut alors calculer la probabilité lissée sur l'ensemble de l'échantillon. Elle est donnée par :

$$P(s_t | \mathcal{Y}_T; \theta) = \sum_{s_T} P(s_T, s_t | \mathcal{Y}_T; \theta)$$

La représentation de ces probabilités en fonction du temps permet de déduire les périodes associées à chaque régime. En effet, si la probabilité d'être dans un état est très proche de 1 pendant une période donnée, on peut en déduire que ces observations ont généré ce régime. En revanche, dès que les probabilités d'être dans cet état tendent vers 0, les observations correspondent à l'autre état (dans le cas de deux états). Les deux types de probabilités sont instructives. Comme leur nom l'indique, les probabilités “lissées” fluctuent moins que les probabilités “filtrées”. En effet, ces dernières varient dès qu'une observation s'écarte un peu de la dynamique correspondant à un régime, puisqu'elles sont calculées à partir de l'information passée uniquement. En revanche, les probabilités “lissées” prennent en compte toute l'information disponible apportée par l'échantillon. Ainsi, elles ne varient que quand le changement de régime est significatif.

³⁰Le filtre utilisé pour calculer les probabilités d'être dans un régime à une date donnée est décrit plus en détails dans Hamilton [1989]. Il est très proche du filtre de Kalman.

5 Conclusion

Nous avons présenté dans cet article les méthodes économétriques récentes relevant des séries temporelles non-linéaires. Plus particulièrement, nous avons présenté trois classes de modèles à changement de régime, tout particulièrement adaptés à expliquer l'évolution des séries économiques, très souvent caractérisées par des asymétries cycliques ou des ruptures dans leur évolution. L'objectif était essentiellement de fournir les bases afin de permettre une utilisation de ces modèles dans le monde économique. Ainsi, nous avons privilégié une présentation pédagogique des méthodes de spécification de ces modèles.

References

- Andrews, D.W.K. et W. Ploberger, Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, Novembre 1994, 62 (6), 1383–1414.
- Beaudry, P. et G. Koop, Do Recessions Permanently Change Output, *Journal of Monetary Economics*, 1993, 31, 149–163.
- Bec, F., M. Ben Salem, et C. Collard, Asymmetries in Monetary Policy Reaction Function : Evidence for the U.S., French and German Central Banks, *Studies in Nonlinear Dynamics and Econometrics*, 2002, 6.
- , — , et R. MacDonald, *Real Exchange Rates and Real Interest Rates : A Non-linear Perspective*, Research Discussion Paper 12, Centre for Financial Markets, University of Strathclyde 1999.
- BenSalem, M. et C. Perraudin, Tests de linéarité, spécification et estimation des modèles à seuil à transition brutale : Une analyse comparée des méthodes de Tsay et de Hansen, *Economie et Prévision*, 2001, 142.
- Caner, M. et B. Hansen, *Threshold Autoregression with a Unit Root*, Miméo, Koc University, Istanbul 1997.
- Chan, K. S. et H. Tong, On estimating thresholds in autoregressive models, *Journal of Time Series Analysis*, 1986, 7, 179–190.
- Davies, R. B., Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika*, 1977, 64, 247–254.
- Eitrheim, Ø. et T. Teräsvirta, Testing the Adequacy of Smooth Transition Autoregressive Models, *Journal of Econometrics*, 1996, 74, 59–75.
- Ertel, J.E. et E.B. Fowlkes, Some Algorithms for Linear Spline and Piecewise Multiple Linear Regression, *Journal of American Statistical Association*, 1976, 71 (355), 640–648.

- Filardo, A.J., Business-Cycle Phases and Their Transitional Dynamics, *Journal of Business & Economic Statistics*, 1994, 12 (3), 299–308.
- Garcia, R. et H. Schaller, *Are the Effects of Monetary Policy Asymmetric?*, Miméo August 1995.
- Godfrey, L. G., *Misspecification Tests in Econometrics*, Cambridge University Press, 1988.
- Goodwin, T. H., Business-Cycle Analysis with a Markov-Switching Model, *Journal of Business & Economic Statistics*, 1993, 11 (3), 331–339.
- Granger, C.W.J. et T. Teräsvirta, *Modelling Nonlinear Economic Relationships*, Oxford University Press, 1993.
- Guégan, D., *Séries chronologiques non-linéaires à temps discret*, Paris: Economica, 1994.
- Haggan, V. et T. Ozaki, Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series models, *Biometrika*, 1981, 68, 189–196.
- Hamilton, J. D., A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 1989, 57, 357–384.
- , Analysis of Time Series Subject to Changes in Regime, *Journal of Econometrics*, 1990, 45, 39–70.
- Hamilton, J.D., Estimation, Inference and Forecasting of Time Series Subject to Changes in Regime, in Rao Maddala et Vinod, éditeurs, *Handbook of Statistics*, Vol. 11, North-Holland, 1993, chapter 9.
- , *Time Series Analysis*, Princeton: Princeton University Press, 1994a.
- Hansen, B.E., Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica*, march 1996a, 64 (2), 413–430.
- , *Sample Splitting and Threshold Estimation*, Mimeo, Boston College April 1996b.
- Hénin, P.Y. et N'Diaye P., *Une Etude des Interactions entre Régimes Budgétaires et Phases Conjoncturelles*, Miméo, Cepremap 1998.
- Karamé, F. et C. Perraudin, *A Comparison of Asymmetric Dynamics of Job Creation and Job Destruction Flows in France and in the United-States*, Miméo, Université Paris 1 1999.
- Keenan, D.M., A Tukey Non-Additivity-Type Test for Time Series Nonlinearity, *Biometrika*, 1985, 72, 39–44.

- Kim, C. J., Dynamic Linear Models with Markov-Switching, *Journal of Econometrics*, 1994, *60*, 1–22.
- Kugler, P., The Term Structure of Interest Rates and Regime Shifts: Some Empirical Results, *Economics Letters*, 1996, *50*, 121–126.
- Ljung, G. et G.E. Box, On a Measure of Lack of Fit in Time Series Models, *Biometrika*, 1978, *65*, 297–303.
- Luukkonen, R., P. Saikkonen, et T. Teräsvirta, Testing Linearity in Univariate Time Series Models, *Scandinavian Journal of Statistics*, 1988, *15*, 161–175.
- , — , et T. Teräsvirta, Testing Linearity against Smooth Transition Autoregressive Models, *Biometrika*, 1988a, *75* (3), 491–499.
- McLeod, A.I. et W.K. Li, Diagnostic Checking ARMA Time Series Models Using Squared Residual Autocorrelations, *Journal of Time Series Analysis*, 1983, *4*, 269–273.
- Neftçi, S. N., Are Economic Time Series Asymmetric Over the Business Cycle?, *Journal of Political Economy*, April 1984, *92* (2), 307–328.
- Ozaki, T., *Non-linear models for non-linear random vibrations*, Technical Report 92, Dept. of Mathematics, University of Manchester, Oxford University 1978.
- Perraudin, C., La dynamique asymétrique de l'emploi au cours du cycle, *Economie et Prévision*, 1995, *4*.
- , *Asymétries conjoncturelles et dynamique de l'emploi : essai de modélisations non-linéaires*, Thèse de Doctorat, Université de Paris I 1999.
- Perron, P., The great crash, the oil price shock and the unit root hypothesis, *Econometrica*, Novembre 1989, *57* (6).
- Petrucelli, J. D., A Comparison of Tests for SETAR-type Non-linearity in Time Series, *Journal of Forecasting*, 1990, *9*, 25–36.
- et N. Davies, A Portmanteau test for Self-exiting Threshold autoregressive, *Biometrika*, 1986, *73*, 687–694.
- Phillips, K. L., A Two-Country Model of Stochastic Output with Change in Regime, *Journal of International Economics*, 1991, *31*, 121–142.
- Potter, S.M., A Nonlinear Approach to US GNP, *Journal of Applied Econometrics*, 1995, *10*, 109–125.
- Press, W.H., S.A. Teukolsky, W.T. Vetterling, et B.P. Flannery, *Numerical Recipes in C*, United States: Cambridge University Press, 1992.

- Priestley, M.B., *Non-Linear and Non-Stationary Time Series Analysis*, London: Academic Press, 1989.
- Ramsey, J.B., Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis, *Journal of the Royal Statistical Society B*, 1969, *31*, 350–71.
- Saikkonen, P. et R Luukkonen, Lagrange Multiplier Tests for Testing Nonlinearities in Time Series Models, *Scandinavian Journal of Statistic*, 1988, *15*, 15–68.
- Schlittgen, R., *Fitting of Threshold Models for Time Series*, Miméo, Institut for Statistik and Econometry, University of Hamburg 1997.
- Teräsvirta, T., Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models, *Journal of the American Statistical Association*, 1994, *89*, 208–218.
- , Modelling Nonlinearity in US Gross National Product 1889-1987, *Empirical Economics*, 1995, *20*, 577–597.
- et H. M. Anderson, Characterizing Nonlinearities in Business Cycles using Smooth Transition Autoregressive Models, in *Nonlinear Dynamics, Chaos and Econometrics*, Pesaran M. H. and Potter S. M., 1993, pp. 111–128.
- Tong, H., *Nonlinear Time Series : A Dynamical System Approach*, Oxford University Press, 1990.
- et K. S. Lim, Threshold Autoregression, Limit Cycles and Data, *Journal of the Royal Statistical Society, Serie B*, 1980, *42*, 245–292.
- Tsay, R., Testing and Modelling Threshold Autoregressive Processes, *Journal of the American Statistical Association*, 1989, *84*, 231–240.
- Warne, A., *Estimation of Vector Autoregressive Models with Changes in Regime via the EM Algorithm*, Gauss Codes, IIES Stockholm University 1995.