# Introduction aux modèles espace état et au filtre de Kalman

Matthieu Lemoine, Florian Pelgrin

▶ **To cite this version:**

Matthieu Lemoine, Florian Pelgrin. Introduction aux modèles espace état et au filtre de Kalman. Revue de l'OFCE, 2003, 86, pp.203-229. 10.3917/reof.086.0203 . hal-01019094

**HAL Id: hal-01019094**

**https://sciencespo.hal.science/hal-01019094**

Submitted on 7 Jul 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# INTRODUCTION AUX MODÈLES ESPACE-ÉTAT ET AU FILTRE DE KALMAN

**Matthieu Lemoine ***
*Département analyse et prévision de l'OFCE*

**Florian Pelgrin ***
*Banque du Canada, Eurequa, Université Paris I et OFCE*

*Nous détaillons ici les principaux concepts et problèmes liés aux modèles espace-état, ainsi que leurs applications. Nous présentons d'abord ces modèles dans leur généralité. Ensuite, nous explicitons les algorithmes utilisés afin de procéder à l'estimation par le maximum de vraisemblance, c'est-à-dire fondamentalement le filtre de Kalman et l'algorithme EM. Nous considérons enfin quatre applications : les décompositions tendance-cycle, l'extraction d'indicateurs coïncidents d'activité, l'estimation d'un taux de chômage d'équilibre pouvant varier avec le temps (TV-Nairu) et l'évaluation du contenu informatif de la courbe des taux sur l'inflation future.*

L'émergence des modèles dynamiques à facteurs ou à variables cachées est relativement récente dans la recherche empirique. Les modèles dynamiques à facteurs linéaires ou modèles espace-état en constituent une classe particulière. Comme cela sera montré dans cet article, de nombreuses procédures statistiques fréquemment utilisées dans la branche empirique de la recherche économique peuvent se reformuler dans le cadre des modèles espace-état, notamment les modèles ARIMA, les modèles à composantes inobservables, les modèles à tendance stochastique, les modèles d'indices coïncidents et les modèles à coefficients aléatoires. Si l'on ne peut pas attendre de ces modèles de déboucher sur des diagnostics *économiques* plus riches que

* Les auteurs tiennent à remercier Eurostat pour leur avoir permis de travailler sur ces sujets dans le cadre du projet « Indicateurs conjoncturels de la zone euro », Mme Boone pour les avoir autorisés à répliquer ses résultats et ses suggestions, ainsi que les participants du séminaire interne de l'OFCE. Ce travail a largement bénéficié des enseignements de M. Monfort à l'Ensae. F. Pelgrin remercie tout particulièrement M. Trognon pour ses conseils dans le cadre d'un enseignement à l'Isup. Selon la formule consacrée, les erreurs et omissions restantes sont de la seule responsabilité des auteurs.

les procédures dont ils sont issus, ce cadre formel présente l'intérêt *statistique* de déboucher sur des estimateurs du maximum de vraisemblance de variables cachées et de fournir les intervalles de confiance associés.

Ce texte a pour objet, d'une part, de détailler et de rappeler les principaux résultats des modèles espace-état et du filtre de Kalman et, d'autre part, de montrer comment ces techniques statistiques peuvent être utilisées face à diverses problématiques économiques. Dans une première section, les modèles espace-état sont présentés avec les algorithmes permettant d'en estimer les variables cachées et les paramètres, notamment le filtre de Kalman et l'algorithme dit EM. Dans une seconde section, quatre applications de cette méthodologie sont décrites : la décomposition en tendance et cycle du PIB américain, la construction d'un indicateur coïncident de l'activité, la détermination du taux de chômage d'équilibre et l'évaluation au cours du temps du contenu informatif de la courbe des taux sur l'inflation future.

# Introduction aux modèles espace-état

L'étude de systèmes physiques émettant au cours du temps des signaux déterminés par des états internes non observés, a conduit à développer en traitement du signal (une branche des sciences de l'ingénieur) les modèles dits espace-état. Dans un premier temps, ces modèles sont présentés en comparaison des techniques économétriques plus usuelles pour modéliser les séries temporelles. Les méthodes d'estimation de tels modèles sont ensuite expliquées en deux temps : l'estimation des variables cachées d'abord (avec le filtre de Kalman), puis celle des paramètres (avec l'algorithme EM) [1].

## Présentation générale des modèles espace-état

Les modèles espace-état intègrent la distinction entre les variables observées (le signal) et les variables cachées (l'état interne). Ils sont constitués :

— d'une ou plusieurs équation(s) de mesure (équation (1) dans l'encadré) décrivant la manière dont les variables observées sont générées par les variables cachées et les résidus.

1. Des présentations plus détaillées des modèles espace-état et des méthodes d'estimation associées sont proposées dans Gourierroux et Montfort (1990) et, avec des extensions aux cas non-linéaires non gaussiens, dans Durbin et Koopman (2001).

— d'une ou plusieurs équation(s) d'état (équation (2) dans l'encadré) décrivant la manière dont les variables cachées sont générées à partir de leur retard et d'innovations.

Les modèles espace-état peuvent s'écrire sous une forme dite développée (encadré) qui implique, sous certaines hypothèses, que la variable de mesure s'écrit sous la forme d'une moyenne mobile. Où réside donc l'intérêt de recourir aux modèles espace-état pour décomposer ou prévoir une série temporelle relativement aux techniques classiques développées par Box et Jenkins (1976) ? D'une part, on s'épargne ici les problèmes de stationnarité [2] et de racine unitaire qui se posent préalablement à l'estimation d'un modèle ARMA. En particulier, les résultats du filtre de Kalman restent valides en présence de séries non-stationnaires. Par ailleurs, ce cadre permet également de relâcher l'hypothèse d'une distribution gaussienne pour les bruits. D'autre part, l'estimation optimale prend en compte l'information disponible à partir de la date initiale $t = 0$, alors que les estimateurs optimaux ARMA prennent en compte l'information à partir de $t = -\infty$ (en pratique, ils sont tronqués à partir de la première date connue). Finalement, les coefficients du modèle peuvent évoluer et ne sont pas obligatoirement considérés comme invariants au cours de la période d'estimation.

Dans leur version élémentaire, les modèles espace-état reposent sur un certain nombre d'hypothèses principales : les équations de mesure et d'état sont linéaires ; les bruits d'observation et d'innovation sont des bruits blancs [3] ; les variables cachées suivent à un instant initial donné une loi gaussienne. À ces dernières, se sont ajoutées des hypothèses secondaires permettant de déterminer la forme canonique (*cf* la définition 2 dans l'encadré) : l'indépendance entre les bruits d'observation et d'innovation (condition d'inversibilité [4]) et l'indépendance entre la variable cachée initiale et ces bruits (condition de causalité). Toutes ces hypothèses sont destinées à simplifier les procédures d'estimation.

Pour autant, on ne peut pas associer à un processus donné $Y_t$ une unique représentation espace-état. En effet, s'il existe une représentation de vecteur d'état $Z_t$, on peut formuler facilement une autre représentation $Z_t^* = E_t Z_t$, $E_t$ étant une matrice inversible quelconque. De même, au lieu de modéliser $Z_{t+1}$ dans l'équation d'état, on pourrait sans difficulté adapter l'estimation à un modèle d'état de $Z_t$. Enfin, diverses dimensions du vecteur d'état sont possibles et il convient de rechercher un modèle de dimension minimale (encadré), de manière à ne pas alourdir la procédure d'estimation.

2. On considère ici la notion de stationnarité au sens faible, c'est-à-dire, pour une série $X_t$, le cas où l'espérance mathématique $E(X_t)$, la variance $V(X_t)$ et les auto-covariances $Cov(X_t, X_{t+h})$ sont indépendantes du temps.

3. Un bruit blanc (au sens faible) est un processus aléatoire d'espérance et d'auto-covariances nulles, dont la distribution n'est pas toujours supposée gaussienne.

4. La condition d'inversibilité n'est pas nécessaire. Lorsque les innovations de l'état et des observations ($\varepsilon_t$) et ($\eta_t$) sont corrélées, on réécrit le modèle sous forme canonique avec des innovations de l'état notées $\xi_t = \varepsilon_t - S_t R^{-1}_t \eta_t$.

## Définitions et notations

La terminologie suivante est utilisée :

| | |
|---|---|
| $Y_t$ | est appelé observation ou variable de mesure ; |
| $Z_t$ | est la variable d'état à la date $t$ ; |
| $\varepsilon_t$ | est le vecteur des innovations à la date $t$ ; |
| $\eta_t$ | est le vecteur des erreurs de mesures à la date $t$ ; |
| $A_t$ | est la matrice de transition ; |
| $C_t$ | est la matrice de mesure ; |
| $X_{1,t}$, $X_{2,t}$ | sont des variables exogènes, prédéterminées ; |
| $C_t Z_t$ | est le signal à la date $t$. |

Soit un processus multidimensionnel $Y_t$, on appelle ***modèle espace-état*** de ce processus, le système (I) décrit par les équations matricielles (1) et (2) :

$$\begin{cases} (1): Z_{t+1} = A_t Z_t + B_t X_{1,t} + \varepsilon_t \\ (2): Y_t = C_t Z_t + D_t X_{2,t} + \eta_t \end{cases} \text{où} \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim NID\left(0, \begin{pmatrix} Q_t & S_t \\ S_t' & R_t \end{pmatrix}\right) \qquad \text{(I)}$$

où les matrices $A_t$, $C_t$ sont de taille $K$x$K$ et $n$x$K$, $B_t$ et $D_t$ sont des matrices déterministes de taille $K_1$x$K$ et $K_2$x$K$ et $Z_0$ est un vecteur aléatoire de loi $N(m,P)$ indépendant du bruit blanc normal.

Le système (I) peut s'écrire sous la ***forme développée*** suivante :

$$\begin{cases} Z_{t+1} = \prod_{j=0}^{t} A_{t-j} Z_0 + \sum_{j=1}^{t} \left( A_{t-j} \prod_{k=0}^{j-1} B_{t-k} \right) X_{1,t-j} + \sum_{j=1}^{t} \left( \prod_{k=0}^{j-1} A_{t-k} \right) \varepsilon_{t-j} + \varepsilon_t \\ Y_t = C_t \left[ \prod_{j=0}^{t} A_{t-j} Z_0 + B_t X_{1,t} + \sum_{j=1}^{t} \left( A_{t-j} \prod_{k=0}^{j-1} B_{t-k} \right) X_{1,t-j} + \sum_{j=1}^{t} \left( \prod_{k=0}^{j-1} A_{t-k} \right) \varepsilon_{t-j} + \varepsilon_t \right] + D_t X_{2,t} + \eta_t \end{cases}$$

Les variables d'état et de mesure s'écrivent donc en fonction de la variable d'état initiale, du passé des erreurs de mesure et des innovations ainsi que des variables exogènes. Cette forme est particulièrement utile lorsqu'on s'intéresse à l'estimateur des moindres carrés généralisés du vecteur d'état ou à l'initialisation du filtre de Kalman.

Le système (I) est dit sous ***forme canonique*** si et seulement si :

$$E(\varepsilon_t \eta_s) = E(\varepsilon_t Z_0) = E(\eta_t Z_0) = 0 \quad \forall t,s=1,\ldots,T$$

Le modèle espace-état est alors dit ***causal*** et ***inversible***.

On appelle ***dimension minimale*** d'un système admettant une représentation espace-état, la plus petite dimension possible du vecteur d'état, $K^*$. En particulier, la représentation est dite minimale si $A_t$ est de taille ($K^*$x$K^*$).

Enfin, le modèle espace d'états (I) est dit ***invariant par rapport au temps*** si les matrices $A_t$, $B_t$, $C_t$ et $D_t$ ne dépendent pas de $t$.

# Estimation des variables d'état par le filtre de Kalman

Afin de simplifier l'analyse, les paramètres [5] définissant le modèle espace-état sont supposés connus et le modèle est supposé dépourvu de variables exogènes. La question consiste alors à estimer à chaque instant $t$ les variables cachées (le vecteur d'état) conditionnellement aux variables observées jusqu'à la date $t$ (le vecteur de mesure).

## Présentation de l'algorithme

Pour calculer des estimations filtrées du vecteur d'état, l'algorithme optimal [6], appelé filtre de Kalman, est utilisé. L'algorithme est structuré en deux étapes reprises d'itération en itération. Les deux premières équations (1 et 2) sont des équations de « mises à jour des mesures » (actualisation) et les deux suivantes (3 et 4) de « mise à jour du temps » (prévision). La première étape concerne les lois de probabilité *a posteriori* qui tiennent compte de l'information à la date t. La seconde étape, à la différence de la première, ne dépend pas des observations à la date $t$ : le calcul peut être fait « hors-ligne », c'est-à-dire sans utiliser les signaux $Y_t$. Enfin, la dernière équation (5) actualise la matrice de gain [7] $K_t$ qui intervient dans les équations précédentes. Chaque itération se résume par les cinq équations suivantes :

$$\begin{cases} (1) : Z^*_{t,t} = Z^*_{t-1,t} + K_t(Y_t - C_t Z^*_{t-1,t}) \\ (2) : \Sigma_{t,t} = (I - K_t C_t)\Sigma_{t-1,t} \\ (3) : Z^*_{t,t+1} = A_t Z^*_{t,t} \\ (4) : \Sigma_{t,t+1} = A_t \Sigma_{t,t} A_t^{'} + Q_t \\ (5) : K_t = \Sigma_{t-1,t} C_t^{'} (C_t \Sigma_{t-1,t} C_t^{'} + R_t)^{-1} \end{cases}$$

et par l'initialisation : $Z^*_{-1,0} = m, \Sigma_{-1,0} = P$

où

— $Z^*_{t,t}$ est l'estimation courante du vecteur d'état ;

— $\Sigma_{t,t} = V(Z_{t,t} - Z^*_{t,t})$ est l'erreur quadratique moyenne sur $Z_t$ ;

5. Il s'agit ici principalement des matrices $A_t$, $C_t$, $R_t$, $Q_t$ et $P$ et du vecteur $m$. Pour simplifier, on écarte ici le cas des variables exogènes, ce qui revient à supposer que $B_t$ et $D_t$ sont nulles.

6. Sous le terme « meilleure approximation » ou « optimal », on pense ici deux critères d'optimalité qui s'avèrent être équivalents dans le cas gaussien : la maximisation de la vraisemblance du vecteur d'état conditionnellement au vecteur de mesure ou la minimisation des carrés des erreurs réalisées sur le vecteur d'état. Dans le cas non-gaussien, le filtre de Kalman reste uniquement optimal parmi les estimateurs linéaires.

7. La matrice $K_t$ est dénommée matrice de gain car, comme cela sera expliqué plus loin, sa prise en compte dans l'équation (2) engendre un gain en précision de l'estimation $Z^*_{t,t}$ de la variable cachée, relativement à $Z^*_{t-1,t}$.

— $Z^*_{t-1,t}$ est la prévision du vecteur d'état faite à la date $t-1$ ;

— $\Sigma_{t-1,t} = V(Z_{t-1,t} - Z^*_{t-1,t})$ est l'erreur quadratique moyenne de prévision correspondante ;

— $K_t$ est la matrice de gain de Kalman qui sera définie dans l'algorithme qui suit.

L'équation (1) calcule l'estimation courante du vecteur d'état $Z^*_{t,t}$ comme la somme pondérée de la prévision à la date $t-1$ du vecteur d'état $Z_t$ et de l'erreur de prévision calculée à partir de la dernière valeur observée $Y_t$. La pondération $K_t$, appelée matrice de gain, est actualisée à chaque itération par l'équation (5). L'équation (3) permet de calculer la prévision de $Z_t$ à la date $t+1$, $Z^*_{t,t+1}$, comme la projection de $Z_{t,t+1}$ sur son passé (passé synthétisé par $Z_{t,t}$).

Les équations (2) et (4) sur les matrices de covariance sont appelées « équations de Riccati ». Ces équations permettent de calculer la suite des gains de Kalman $K_t$ et ce calcul peut être fait « hors-ligne ». La matrice de covariance *a posteriori* $\Sigma_{t,t}$ connaît généralement un gain en précision par rapport à la matrice de covariance *a priori* $\Sigma_{t-1,t}$ grâce au terme $K_t C_t \Sigma_{t-1,t}$ (équation 2). La matrice de covariance *a priori* en $t+1$, $\Sigma_{t,t+1}$, prend en compte les erreurs liées aux innovations de l'état avec la matrice $Q_t$, mais est aussi augmentée d'un terme $A_t \Sigma_{t,t} A'_t$ associé aux erreurs sur l'état à la date $t$ (équation 4). Lorsque les variables d'état sont stationnaires, la covariance prévue $\Sigma_{t,t+1}$ qui part d'une incertitude *a priori* $P$, tend vers une constante $\Sigma_\infty$ (voir Harvey, 1989). Après une période transitoire, les intervalles de confiance entourant des variables cachées stationnaires ont donc une largeur à peu près constante.

## Extensions à d'autres modes d'estimation

L'estimation du vecteur d'état envisagée jusqu'ici est une estimation filtrée et se distingue des estimations lissées ou prévues :

— le filtrage consiste à rechercher la meilleure approximation de l'état $Z_t$ sachant les observations présentes et passées $Y_0, ..., Y_t$

— la prévision consiste à rechercher la meilleure approximation de l'état $Z_t$ sachant les observations passées $Y_0, ..., Y_{t-h}$

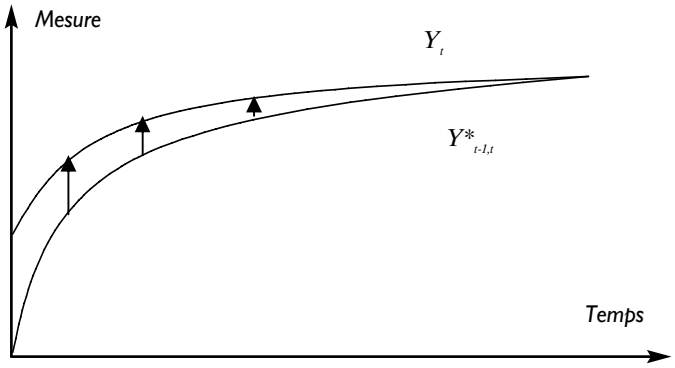
— le lissage consiste à rechercher la meilleure approximation de l'état $Z_t$ sachant les observations passées, présentes et futures $Y_0, ..., Y_T$

Les problèmes de prévisions et de lissage se traitent à partir de simples extensions de l'algorithme de filtrage présenté précédemment.

## Interprétations graphiques du filtre de Kalman

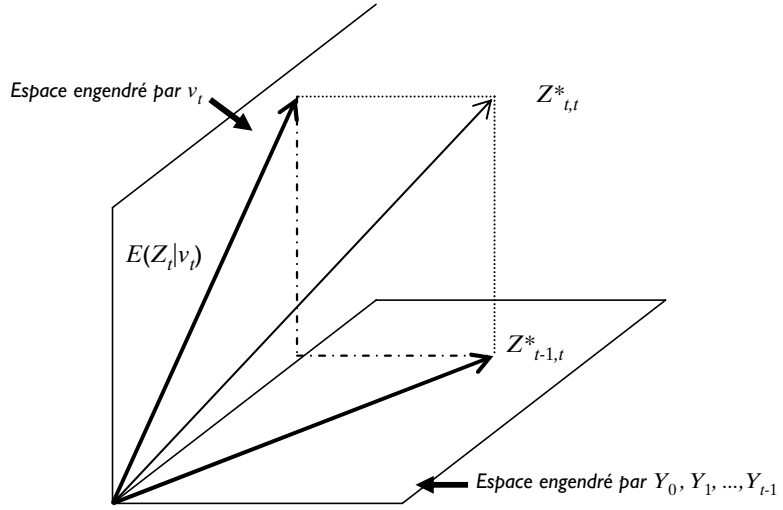
Si le filtre de Kalman vise d'abord à estimer la variable cachée $Z_t$, il permet également de réaliser une prévision de $Y_t$, connaissant son passé. Il suffit pour cela de calculer $Y^*_{t-1,t} = C_t Z^*_{t-1,t}$. En raisonnant sur la variable d'observation, le filtre calcule en fait une prévision $Y^*_{t-1,t}$ qui est ensuite corrigée par comparaison avec la mesure réelle $Y_t$. Le filtre fait donc dévier progressivement la trajectoire théorique $Y^*_{t-1,t}$ (engendrée par les conditions initiales et l'équation d'état) vers la trajectoire réelle $Y_t$, comme le montre le schéma 1. En raisonnant sur l'état $Z_t$, son estimation actualisée $Z^*_{t,t}$ est décomposable entre la prévision $Z^*_{t-1,t}$ et un terme correctif proportionnel à l'erreur relative aux dernières observations $Y_t$.

Schéma 1 : Trajectoire théorique et trajectoire réelle




Une seconde représentation du filtre de Kalman consiste à utiliser des propriétés élémentaires de l'algèbre linéaire. Rechercher $Z^*_{t,t} = E(Z_t|Y_{1,t})$ revient en effet à projeter $Z_t$ sur l'espace engendré par $Y_1$, …, $Y_t$. Or cet espace peut se décomposer en l'espace des valeurs observées par le passé (engendré par $Y_1$, …, $Y_{t-1}$) et celui de l'innovation $v_t$ (égale à $Y_t - C_t Z^*_{t-1,t}$). En projetant $Z_t$ sur ces deux sous-espaces, on obtient par définition $Z^*_{t-1,t}$ et $E(Z_t|v_t)$. On peut alors montrer que $E(Z_t|v_t)$ peut se réécrire $K_t v_t$ avec la matrice de gain calculée dans l'équation (5). L'équation (1) exprime alors la projection recherchée comme la somme des projections sur chaque sous-espace (schéma 2).

Schéma 2 : Projection du vecteur $Z^*_{t,t}$




# Estimation des paramètres par le maximum de vraisemblance

## Algorithme EM

Dans la partie précédente, les matrices $A_t$, $C_t$, $Q_t$, $R_t$, $P$ ainsi que le vecteur $m$ étaient supposés connus. En pratique, ces matrices sont inconnues et doivent être estimées. L'algorithme EM est couramment utilisé pour déterminer les Estimateurs du Maximum de Vraisemblance (EMV) des paramètres d'un modèle espace-état. Cet algorithme itératif a le mérite d'être simple, même s'il est relativement lent à converger par rapport à des algorithmes plus sophistiqués. Il a été introduit par Dempster *et al.* (1977) pour estimer le maximum de vraisemblance de modèles stochastiques à variables cachées.

Pour procéder à une estimation par maximum de vraisemblance des paramètres d'un modèle espace-état, il est nécessaire d'avoir l'expression de la fonction de vraisemblance. Pour chaque jeu de paramètres $\theta$, la log-vraisemblance associée à un échantillon $Y_1,...,Y_T$ d'un modèle espace-état s'exprime à partir des valeurs prévues de l'état $Z^*_{t-1,t}$ et des matrices de covariance associées $\Sigma_{t-1,t}$ :

$$ln\,l\left(Y_{0:T};\theta\right) = cte - \frac{1}{2}\sum_{t=0}^{T} ln\,det\,M_{t-1,t}(\theta) - \frac{1}{2}\sum_{t=0}^{T}\tilde{Y}'_{t-1,t}(\theta)M^{-1}_{t-1,t}(\theta)\tilde{Y}_{t-1,t}(\theta)$$

avec $\tilde{Y}_{t-1,t}(\theta) = Y_t - C_t Z^*_{t-1,t}$ et $M_{t-1,t}(\theta) = C_t \Sigma_{t-1,t} C'_t + R_t$

L'algorithme EM est alors un algorithme itératif qui génère une séquence d'estimations $(\theta_i)_{i=1,2,...}$ à partir d'une condition initiale $\theta_0$. Chaque itération se décompose en deux étapes qui s'écrivent :

— Étape E : $ln\ l(Y_{0:T};\theta)$ se déduit de $Z^*_{t-1,t}(\theta_i)$ et de $\Sigma_{t-1,t}(\theta_i)$, calculés par un filtre de Kalman.

— Étape M : la maximisation de $ln\ l(Y_{0:T};\theta)$ par rapport à $\theta$ conduit à $\Sigma_{i+1}$.

La première étape E (« *Expectation* », c'est-à-dire « Espérance ») calcule une vraisemblance à partir de la formule précédente sur la vraisemblance d'un modèle espace-état. Ces formules mobilisent en particulier l'application d'un filtre de Kalman pour connaître l'espérance conditionnelle de l'état $Z^*_{t-1,t}$ et de sa covariance $\Sigma_{t-1,t}$ à paramètres $\theta_i$ et observations $Y_{0:T}$ fixés. La seconde étape M (« *Maximisation* »), consiste à rechercher un jeu de paramètres maximisant la vraisemblance estimée dans l'étape E. Cette maximisation peut-être analytique ou numérique selon la complexité du problème. Après un cycle « Étape E / Étape M », on obtient $\theta_{i+1}$ et on peut montrer que $L(Y_t|\theta_{i+1})>L(Y_t|\theta_i)$. En itérant ces étapes E et M, les paramètres estimés par l'algorithme convergent généralement vers le maximum de vraisemblance.

## Quelques limites pratiques

Les propriétés statistiques de l'estimateur du maximum de vraisemblance ne sont pas abordées ici mais certaines difficultés de la phase d'estimation sont présentées. Trois problèmes sont brièvement étudiés : le choix des conditions initiales, l'importance du ratio signal/bruit et les propriétés de convergence de l'algorithme EM.

La mise en œuvre du filtre de Kalman nécessite généralement de spécifier les conditions initiales du vecteur d'état. En effet, si tous les éléments du vecteur d'état initial $Z_0$ sont exactement connus *a priori*, alors $Z_0$ a une distribution *a priori* correcte, c'est-à-dire dont tous les moments sont finis, avec une moyenne connue et une matrice de variance-covariance bornée. Le filtre de Kalman fournit alors la fonction de vraisemblance exacte des observations par la décomposition de l'erreur de prévision. Une telle information *a priori* est cependant rarement disponible. Dans cette perspective, une première méthode consiste à fixer arbitrairement les valeurs initiales du vecteur d'état $Z_0$. Le problème est que les estimations vont dépendre de ces valeurs. Il s'agit alors de tester la sensibilité aux conditions initiales et cela d'autant que l'algorithme EM (voir plus loin) fournit des maxima locaux. Une autre solution est d'estimer les éléments de $Z_0$ par la méthode des moindres carrés généralisés en utilisant la formule développée (encadré 1). Il est à noter qu'il existe de nombreuses variantes de ces deux méthodes.

Un deuxième problème concerne le traitement des matrices $Q$ et $S$ (encadré 1), qui représentent respectivement les matrices de variance-covariance du vecteur des innovations et du vecteur des erreurs de mesure. En effet, un élément fondamental dans l'estimation des modèles espace-état est le degré de lissage des variables non observées, qui dépend des deux matrices précédentes (on suppose que $S$ est la matrice nulle). Par exemple, dans le cas univarié, un ratio $R/Q$ élevé (appelé ratio signal/bruit) contribue à accroître le pouvoir explicatif de la variable latente et l'équation de mesure sera donc mieux estimée. À la limite, pour de grandes valeurs de $Q$, la variable non observée absorbe toute la variation des résidus dans l'équation de mesure. Alternativement, si $Q$ est une matrice nulle et si $A$ est la matrice identité, les estimations filtrées (respectivement lissées) correspondront à la méthode des moindres carrés récursifs (respectivement des moindres carrés). Il est donc particulièrement important de déterminer ce ratio. Dans la pratique, la plupart des études fixent ce ratio de telle sorte que l'estimation de la variable latente soit suffisamment lisse, avec des fluctuations jugées raisonnable d'une période à l'autre. Des tests de sensibilité sont alors utilisés en spécifiant différentes valeurs pour ce ratio.

Un troisième problème est lié aux propriétés de la vraisemblance et à la détermination numérique de la solution. Dans de nombreux cas, la vraisemblance possède malheureusement des maxima locaux dans lesquels l'algorithme peut être piégé. Il est alors préférable de fournir à l'algorithme des paramètres initiaux « relativement proches » du maximum global.

D'autres problèmes pratiques sont abordés dans les applications économiques.

# Applications

La mise en œuvre de la méthodologie précédente est décrite ici dans quatre champs d'application majeurs en économie : les tendance et cycle de production, les indices coïncidents d'activité, les taux de chômage d'équilibre pouvant varier avec le temps (*Time Varying Nairu*, dit TV-Nairu) et le contenu informatif de la courbe des taux sur l'inflation future. Après un bref rappel de chaque contexte économique, il s'agit principalement de montrer comment ont été formulés et estimés les modèles espace-état dans ces différents cas. Les applications présentées dans cette partie sont donc uniquement illustratives et n'ont pas vocation à aborder chaque problématique économique de façon exhaustive. Pour aller plus loin, le lecteur peut se référer aux articles dont ces applications sont inspirées, soient Bentoglio, Fayolle et Lemoine (2001), Stock et Watson (1991), Boone (2000) et l'article précurseur de Garbade et Wachtel (1978).

## Décomposition tendance/cycle [8]

Un premier exemple d'application des techniques d'estimation précédentes concerne les décompositions tendance/cycle qui visent à donner une idée du niveau de la croissance potentielle et des phases du cycle d'affaire. Parmi la diversité des méthodes existantes, les modèles à composantes inobservables développés par Harvey (1989) ont été retenus, car ils se mettent facilement sous forme espace-état. Ces modèles de décomposition sont considérés ici à un niveau univarié, ce qui en restreint le contenu économique, mais a l'avantage de présenter une plus grande lisibilité. Leurs résultats ont de ce fait essentiellement un intérêt descriptif sur le passé [9]. Les résultats présentés proviennent d'une application au cas américain sur des données trimestrielles de PIB (en logarithmes) provenant des comptes nationaux sur la période 1960-2002. Tous les résultats sont obtenus à partir du logiciel STAMP développé par Koopman *et al.* (2000).

Dans le modèle proposé par Harvey (1989), une série $Y_t$ est décomposée, de manière additive [10], en une tendance $T_t$, un cycle $C_t$ et une composante irrégulière $\varepsilon_t$. Chacune de ces composantes est stochastique et elles sont supposées mutuellement non corrélées entres elles. La tendance suit un modèle « localement linéaire » [11], un modèle de tendance assez général permettant de prendre en compte divers types de chocs permanents. La composante cyclique est un processus linéaire stationnaire susceptible de faire apparaître une alternance relativement régulière de pics et de creux, tout en admettant une certaine persistance des phases du cycle et d'éventuelles dissymétries entre elles. Sa modélisation requiert l'introduction d'une variable $C^*_t$ qui n'a pas d'interprétation particulière. La composante irrégulière $\varepsilon_t$ est un bruit blanc.

$$\begin{cases} Y_t = T_t + C_t + \varepsilon_t \\ T_t = T_{t-1} + \beta_{t-1} + \eta_t \\ \beta_t = \beta_{t-1} + \zeta_t \\ C_t = \rho\left(cos\,\lambda.C_{t-1} + sin\,\lambda.C^*_{t-1}\right) + \kappa_t \\ C^*_t = \rho\left(-sin\,\lambda.C_{t-1} + cos\,\lambda.C^*_{t-1}\right) + \kappa^*_t \end{cases}$$

8. Pour une présentation plus détaillée, voir Bentoglio, Fayolle et Lemoine (2001).

9. Voir dans cette même Revue l'article « Écart de production dans la zone euro : une estimation par le filtre HPMV » de Chagny et Lemoine (2003), pour une décomposition tendance/cycle appliquée à la zone euro et visant à donner un contenu économique plus riche à ce type d'approche statistique.

10. Pour réaliser une décomposition multiplicative d'une série temporelle, on peut procéder de la même manière en considérant la transformation logarithmique d'une telle série.

11. Pour une présentation détaillée des modèles localement linéaires, on pourra consulter Durbin et Koopman (2001). Ce modèle permet aux aléas d'être localisés sur le niveau ou la pente de la tendance. Les tendances déterministes, I(1) ou I(2) constituent des cas particuliers de ce modèle.

où $\varepsilon_t$, $\eta_t$, $\zeta_t$, $\kappa_t$ et $\kappa^*_t$ sont des bruits blancs indépendants, respectivement de variance $\sigma^2_\eta$ et $\sigma^2_\zeta$, $\sigma^2_\kappa$ et à nouveau $\sigma^2_\kappa$. Le paramètre $\rho$, compris entre $0$ et $1$, est un paramètre de lissage : plus il est proche de $0$, plus le cycle est irrégulier ; plus il est proche de $1$, plus le cycle est lisse. En l'absence de chocs sur le cycle ($\sigma^2_\kappa = 0$) et avec un lissage maximal ($\rho = 1$), le cycle est une sinusoïde parfaite dont la période se déduit du paramètre $\lambda$ par la formule $2\pi/\lambda$. Mais, dans le cas général, les chocs $\kappa_t$ et $\kappa^*_t$ peuvent avancer ou retarder les reprises comme les ralentissements et $2\pi/\lambda$ ne représente plus que la durée moyenne du cycle.

Pour estimer les paramètres ($\sigma^2_\varepsilon$, $\sigma^2_\eta$, $\sigma^2_\zeta$, $\sigma^2_\kappa$, $\lambda$, $\rho$), la tendance et le cycle, on reformule le modèle à composantes inobservables précédent comme un **modèle espace-état**, avec l'*équation d'observation* :

$$Y_t = (1\ 0\ 1\ 0)X_t + \varepsilon_t$$

où $X_t$ est le vecteur d'état défini par $X_t = (T_t\ \beta_t\ C_t\ C^*_t)'$ qui suit l'*équation d'état* :

$$X_t = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & -\rho\sin\lambda & \rho\cos\lambda \end{pmatrix} X_{t-1} + \begin{pmatrix} \eta_t \\ \zeta_t \\ \kappa_t \\ \kappa^*_t \end{pmatrix}$$

L'algorithme EM permet alors d'estimer les paramètres $(\sigma_\varepsilon\ \sigma_\eta\ \sigma_\zeta\ \sigma_\kappa\ \rho\ \lambda)'$ du modèle. Pour ce faire, le vecteur d'initialisation est fixé à $(10^{-4}\ 10^{-4}\ 10^{-4}\ 10^{-4}\ 0{,}2\ 0{,}8)'$. La vraisemblance est alors maximisée avec l'algorithme EM et on obtient les valeurs estimées présentées avec leur écart-type dans le tableau 1. Les variances $\sigma_\varepsilon$ et $\sigma_\eta$ de la composante irrégulière et du niveau de la tendance sont estimés à $0$ et la vraisemblance est concentrée par rapport à ceux-ci. Ensuite, un filtre de Kalman fournit les innovations $v_t$ de $Y_t$, leur variance $F$ et la matrice de gain $K_t$. Un lisseur appliqué aux différents bruits $(\zeta, \kappa, \kappa^*, \varepsilon)$ permet *in fine* de générer la tendance estimée $\underline{T}_t$ et le cycle estimé $\underline{C}_t$.

1. Paramètres de la décomposition du PIB américain

| | $\sigma_\varepsilon$ | $\sigma_\eta$ | $\sigma_\zeta$ | $\sigma_\kappa$ | $\lambda$ | $\rho$ |
|---|---|---|---|---|---|---|
| Valeur estimée | 0 | 0 | 0,04 | 0,75 | 0,20 | 0,95 |
| (écart-type) | (-) | (-) | (0,01) | (0,08) | (0,02) | (0,00) |

*Note* : les paramètres $\sigma_\varepsilon$ $\sigma_\eta$ $\sigma_\zeta$ $\sigma_\kappa$ et leur écart-type sont multipliés par 100.
*Sources* : Bureau of Economic Analysis, calculs des auteurs.

2. Statistiques des erreurs de prévision

| Statistiques | $\sigma$ | $N$ | $Q(13,10)$ |
|---|---|---|---|
| Résidus | $8{,}5.10^{-3}$ | 5,2 | 14,0 |

*Note* : Pour tester la normalité des résidus, est calculée la statistique de Bowman-Shenton N qui suit un $\chi^2$ à 2 degrés de liberté et admet donc pour valeur critique 6,0 au seuil de 5%. Pour tester la nullité de l'auto-corrélation des p=13 premiers résidus d'un modèle à $n$=4 paramètres, est calculée la statistique de Box-Ljung $Q(p,p\text{-}n\text{+}1)$ qui suit un $\chi^2$ à $p\text{-}n\text{+}1=10$ degrés de liberté et admet donc ici la valeur critique 18,3 au seuil de 5 %.

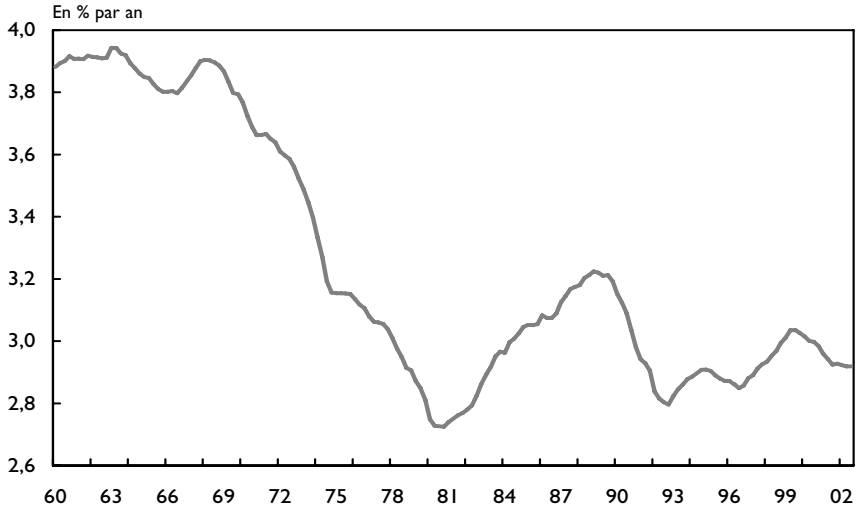
*Sources* : Bureau of Economic Analysis, calculs des auteurs.

Comme le montre le tableau 1, le modèle univarié adéquat est spontanément un modèle du type « tendance douce » [12], c'est-à-dire que les innovations sur le niveau de la tendance ont une variance $\sigma_\eta$ estimée à 0. Le modèle comporte également un cycle stochastique de paramètres $\rho = 0{,}95$ et $1 = 0{,}2$, soit une durée moyenne d'environ 8 ans [13]. La période estimée du cycle américain est raisonnable, mais apparaît plus longue que la durée de 6 ans qui lui est souvent attribuée. Quant au paramètre $\rho$, sa valeur élevée indique que le cycle est relativement lisse, ce qui est conforté par la nullité de la composante irrégulière $(\sigma_\varepsilon = 0)$ : comme l'ont montré Bentoglio *et al.* (2001) en comparaison avec l'Europe, ceci indique que les États-Unis connaissent des périodes d'expansion et de ralentissement régulières, relativement au caractère heurté de la conjoncture européenne. Pour apprécier l'ordre de grandeur des écarts-types des innovations affectant la tendance et le cycle, ainsi que celui de l'erreur de prévision, on peut considérer qu'ils sont exprimés en proportion du niveau moyen de la série. On notera que la variance des innovations affectant la pente de la tendance est bien inférieure à celle des innovations du cycle, ce qui traduit la progressivité des inflexions de la pente de la tendance. Concernant les erreurs de prévision, les statistiques des tests de normalité et de Box-Ljung ne permettent pas de rejeter l'hypothèse nulle selon laquelle il s'agit d'un bruit blanc (tableau 2).

12. Diebold et Rudebusch (1999, p. 20) confirment l'intérêt d'un tel modèle de tendance pour les États-Unis : « For US real output, it appears that a trend representation that is very smooth, even if not exactly linear, is a viable candidate ».

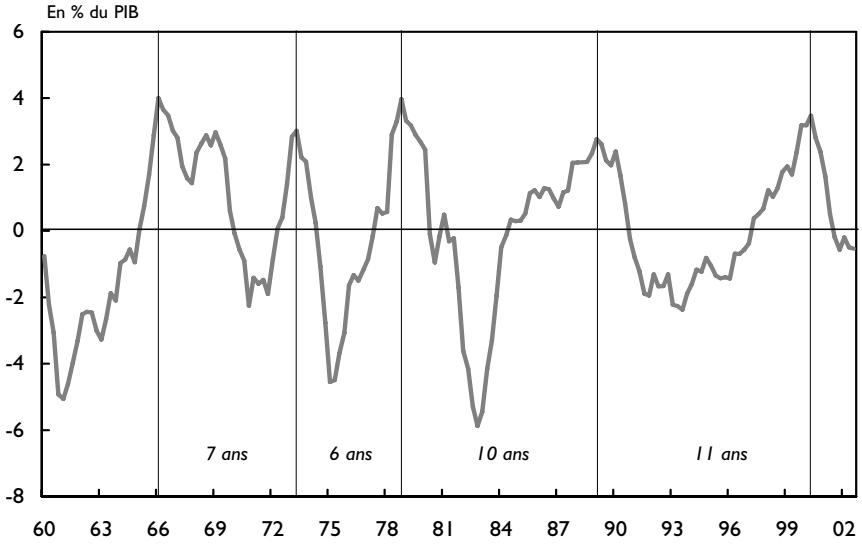
13. Pour mémoire, la durée moyenne du cycle du paramètre $\lambda$ se calcule par la formule $2\pi/\lambda$. Ici, la durée moyenne du cycle vaut donc environ $2*3{,}14/0{,}2$ = 31,4 trimestres, soit environ 8 ans.

1. Taux de croissance de la tendance du PIB américain




*Sources* : Bureau of Economic Analysis, calculs des auteurs.

2. Cycle du PIB américain




*Note* : les barres verticales indiquent les 5 pics cycliques repérés dans la période d'estimation. Sont également ajoutées sur le graphique les durées de ces 5 cycles.
*Sources* : Bureau of Economic Analysis ; calculs des auteurs.

En ce qui concerne la croissance de la tendance, dont l'estimation lissée est présentée sur le graphique 1, elle s'interprète comme un taux de croissance équilibré de long terme qui baisserait de 4,5 % dans les années 1960 à moins de 3 % dans les années 1980 pour revenir à environ 3,5 % à la fin des années 1980, sans compter la part cyclique de l'envolée de la croissance américaine sur cette même période. Le graphique 2 représente l'estimation du cycle américain. On repère alors sur le cycle cinq pics entre 1960 et 2002 : un léger ralentissement initié en 1966 en période de croissance forte, les deux chocs pétroliers en 1973 et en 1979, le pic de 1989 associé à l'éclatement de la bulle immobilière américaine et celui de 2000 à la suite de l'effondrement des valeurs technologiques. Les États-Unis ont donc connu quatre cycles entiers sur la période et on retrouve bien une durée moyenne d'environ 8 ans et non de 6 ans. À l'aide du graphique 2, on voit que cela s'explique par la période retenue : les deux derniers cycles d'une durée très longue (10 ans) prennent une forte importance relative, alors que les cycles étaient plus courts (6 ans) durant les années 1950, 1960 et 1970.

Si ce modèle conduit à d'intéressants résultats sur les évolutions du cycle et de la tendance, sa principale limite concerne l'hypothèse d'indépendance entre la tendance et le cycle. Cette hypothèse exclut tout effet d'hystérèse, notamment l'idée qu'une récession puisse engendrer un ralentissement durable de la tendance. La spécification retenue ici pour le cycle ne permet en fait pas d'identifier la corrélation entre la tendance et le cycle. Comme l'ont montré Morley, Nelson et Zivot (2002), l'identification de cette corrélation requière de modéliser le cycle comme un processus auto-régressif d'ordre 2. Le modèle de Morley *et al.* (2002) n'a pas été retenu ici car, à la différence de $\rho$ et $\lambda$, les paramètres du cycle ne sont pas interprétables et surtout, comme l'a montré Proietti (2002), parce que le cycle issu de ce modèle est sujet à de très importantes révisions avec l'arrivée de nouvelles données.

## Indice coïncident [13]

Nombre de variables macroéconomiques présentent des co-mouvements avec le cycle des affaires. Dans cette perspective, il est utile de construire un indice coïncident à partir de ces variables. Il s'agit alors d'exprimer chaque variable en fonction d'une composante inobservable commune, l'indice coïncident « résumant » les informations de chaque variable économique. Avec cet objectif, le *Departement of Commerce* (DOC) des États-Unis a construit avec une procédure *ad hoc* un indicateur de base 100 en 1995, dénommé DOC. Cet indicateur combine quatre variables macro-économique $(Y_i)$ : l'indice de la production

13. Pour une présentation plus détaillée, voir Stock et Watson (1991).

industrielle $(Y_1)$, le revenu nominal avant transferts $(Y_2)$, les ventes industrielles et commerciales $(Y_3)$ et l'emploi salarié non agricole $(Y_4)$. Pour trouver une solution optimale en terme de vraisemblance à ce problème d'estimation d'un résumé de quatre séries temporelles, Stock et Watson (1991) ont proposé de formuler et d'estimer un modèle espace-état avec des variables cachées appropriées. Nous reprenons ici l'approche de Stock et Watson (1991) et la nouvelle formulation proposée par Kim et Nelson (1999).

Dans un cadre espace-état, les équations de mesure décomposent chaque variation [14] des variables économiques en fonction d'une composante commune [15] d'une part (qui servira à construire l'indice coïncident) et des composantes idiosyncratiques d'autre part. Les équations d'état servent ensuite à modéliser les dynamiques auto-régressives de la composante commune [16] et des composantes idiosyncratiques.

$$(I)\begin{cases} \Delta Y_{it} = D_i + \gamma_i(L)\Delta C_t + e_{it} \\ (\Delta C_t - \delta) = \varphi_1(\Delta C_{t-1} - \delta) + \varphi_2(\Delta C_{t-2} - \delta) + \omega_t, \omega_t \sim N(0,1) \\ e_{it} = \psi_{i1} e_{i,t-1} + \psi_{i2} e_{i,t-2} + \varepsilon_{it}, \varepsilon_{it} \sim N(0, \sigma_i^2) \end{cases}$$

## *Première étape* : Estimation du modèle en écart à la moyenne

Le terme constant $\delta$ dans l'équation de mesure n'étant pas directement identifiable, le modèle est estimé, dans un premier temps, avec un vecteur de mesure $(y_{it})$ contenant les variables différenciées à l'ordre 1 en écart à leur moyenne respective. Quant au vecteur d'état, il contient la composante commune $c_t$, les composantes idiosyncratiques $(e_{it})$ et leurs retards respectifs. Le modèle espace-état précédent se reformule alors par le modèle centré suivant :

$$(II)\begin{cases} \Delta y_{it} = D_i + \gamma_i(L)\Delta C_t + e_{it} \\ \Delta c_t = \varphi_1 \Delta c_{t-1} + \varphi_2 \Delta c_{t-2} + \omega_t, \omega_t \sim N(0,1) \\ e_{it} = \psi_{i1} e_{i,t-1} + \psi_{i2} e_{i,t-2} + \varepsilon_{it}, \varepsilon_{it} \sim N(0, \sigma_i^2) \end{cases}$$

avec $\Delta y_{it} = \Delta Y_{it} - \Delta \overline{Y}_{it}$ et $\Delta c_t = \Delta C_t - \delta$.

*Vecteur d'état :* $Z_t = (\Delta C_t, \Delta C_{t-1}, \Delta e_{it}, \Delta e_{it-1})'_{i=1...4}$

*Vecteur de mesure :* $\Delta y_t = (\Delta Y_{it})'_{i=1...4}$

14. Les variables sont considérées en variation car les tests de racine unitaire (ADF) indiquent qu'elles sont toutes intégrées d'ordre un.

15. Dans l'article original de Stock et Watson (1991), les variable $Y_i$ sont considérées comme coïncidentes à l'activité et ne dépendent dans les équations de mesure que de la valeur courante de la variable $C_t$. Mais les auteurs remarquent que les résidus de l'équation d'emploi $(Y_4)$ restent auto-corrélés. L'emploi $(Y_4)$ est en effet retardé sur l'activité. C'est pourquoi on reprend ici la formulation de Kim et Nelson (1999) d'une équation de mesure spécifique pour l'emploi qui intègre en plus trois retards de $C_t$.

16. La variance des innovations $w_t$ est fixée à 1 pour normaliser la composante commune $C_t$.

Ce modèle est alors estimé et ses paramètres sont reportés dans le tableau 3.

3. Estimation des paramètres

| | $\Delta c_t$ | $\Delta y_{1,t}$ | $\Delta y_{2,t}$ | $\Delta y_{3,t}$ | $\Delta y_{4,t}$ |
|---|---|---|---|---|---|
| $\phi_1$ | 0,5657 | - | - | - | - |
| | *(0,0808)* | | | | |
| $\phi_2$ | 0,0353 | - | - | - | - |
| | *(0,0855)* | | | | |
| $\psi_{i,1}$ | - | – 0,0018 | – 0,0371 | 0,0896 | – 0,0021 |
| | | *(0,0029)* | *(0,0094)* | *(0,0065)* | *(0,0119)* |
| $\psi_{i,2}$ | - | – 0,0338 | – 0,0044 | 0,6811 | 0,0101 |
| | | *(0,0472)* | *(0,0641)* | *(0,0469)* | *(0,0108)* |
| $\sigma^2_i$ | - | 0,1439 | 0,3571 | 0,0152 | 0,1383 |
| | | *(0,0080)* | *(0,0693)* | *(0,0016)* | *(0,0101)* |
| $\gamma_{i,0}$ | - | 0,0647 | 0,2854 | 0,2490 | 0,5762 |
| | | *(0,0384)* | *(0,0487)* | *(0,0253)* | *(0,0334)* |
| $\gamma_{i,1}$ | - | - | - | - | 0,2143 |
| | | | | | *(0,0149)* |
| $\gamma_{i,2}$ | - | - | - | - | 0,4512 |
| | | | | | *(0,0341)* |
| $\gamma_{i,3}$ | - | - | - | - | 0,1383 |
| | | | | | *(0,0101)* |

*Note* : les écart-types des paramètres estimés sont indiqués entre parenthèses.
*Sources* : Bureau of Economic Analysis, Bureau of Labor Statistics, calculs des auteurs.

## *Deuxième étape* : Estimation de l'indicateur coïncident

Dans un second temps, on estime $\delta$ et on construit l'indicateur coïncident $(C_{t,t})$. La forme développée [17] du modèle (I) implique une relation entre $C_{t,t}$ et $\Delta Y_t$ de type :

$$C_{t,t} = W(L)\Delta Y_t \text{ dont on déduit } \hat{\delta} = W(1)\Delta\overline{Y}$$

Le problème se déplace donc à l'identification de $W(1)$. L'astuce proposée par Stock et Watson (1991) consiste alors à identifier $W(1)$ en repassant par le modèle en écart à la moyenne (II). Dans le

17. Pour des détails sur la notion de forme développée d'un modèle espace-état, voir l'encadré.

modèle (II), l'estimateur $c_{t,t}$ de la variable cachée se déduit en effet de la variable observée en utilisant la même forme développée et on obtient :

$$C_{t,t} = W(L)\Delta Y_t$$

Sans rentrer dans les détails, indiquons juste que $W(L)$ et *a fortiori* $W(1)$ peuvent alors se déduire du filtre de Kalman estimé précédemment pour le modèle (II).

Une fois la constante estimée, tous les paramètres du modèle sont connus et l'indicateur coïncident se définit par la formule :

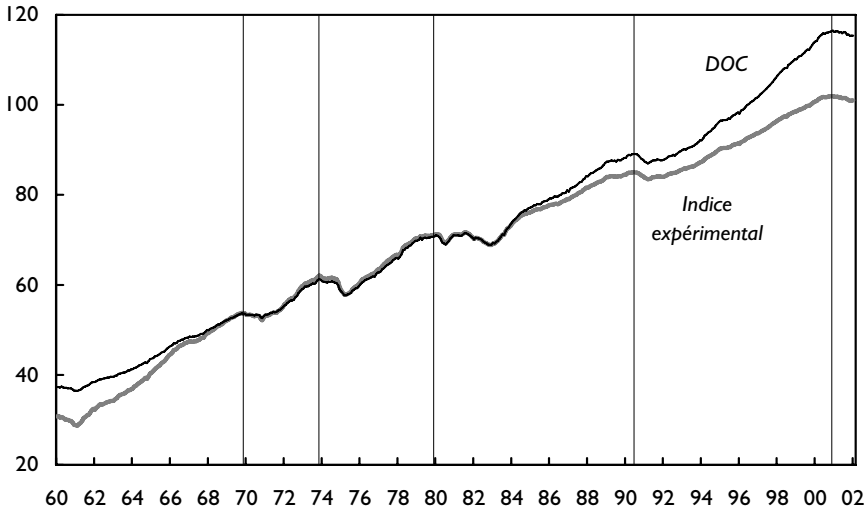
$$C_{t,t} = C_{t,t\text{-}1} + c_{t,t} + \delta$$

L'indicateur est donc identifié au choix arbitraire près d'une valeur initiale $C_{0,0}$. Finalement, on définit un indicateur normalisé (ayant la même valeur en 1970 que l'indice DOC) à partir de la formule suivante :

$$C^*_{t,t} = C_{t,t} + \frac{\sigma_{\Delta C_t^{DOC}}}{\sigma_{\Delta C_{t,t}}}$$

Le graphique 3 montre que les points de retournement du nouvel indicateur et de l'indice DOC sont synchronisés. En revanche, le nouvel indicateur présente une plus forte croissance dans les années 1960 et une croissance plus ralentie dans les années 1980 et 1990.

3. Indicateur coïncident et indice DOC




*Note* : DOC est un indice de base 100 en 1995. L'indice expérimental a un niveau arbitraire égal à l'indice DOC en janvier 1970. Les barres discontinues indiquent les points de retournement de l'indice DOC.
*Sources* : Bureau of Economic Analysis ; Bureau of Labor Statistics ; calculs des auteurs.

## Détermination du taux de chômage d'équilibre [18]

Le Nairu est défini comme le taux de chômage stabilisant l'inflation en l'absence de chocs d'offre temporaires. Si le taux de chômage est inférieur au Nairu, l'inflation s'accroît jusqu'à ce que le taux de chômage atteigne le niveau du Nairu. L'inflation se stabilise alors à un niveau plus élevé. Les évaluations structurelles du Nairu sont réalisées à partir d'une équation de prix et d'une équation de salaires. Ce type d'évaluation pose deux problèmes : la diversité des chiffrages auxquels conduisent les études selon la spécification retenue ; la difficulté à expliquer, pour la fin des années 1990, la compatibilité entre une forte baisse du chômage américain et le maintien d'une inflation modérée.

C'est pourquoi d'autres méthodes d'estimation ont été proposées dans la littérature récente [19] : le filtre Hodrick-Prescott (HP), le filtre Hodrick-Prescott Multivarié (HPMV) [20] ou un filtre de Kalman [21]. Si le filtre HP est un simple lissage du taux de chômage, le filtre HPMV et le filtre de Kalman prennent en compte l'information apportée par l'inflation sur le niveau du Nairu, qui est mesuré comme une variable cachée dans un modèle réduit, le *triangle model* (Gordon, 1997) :

$$\Delta\pi_t = \alpha(L)\Delta\pi_{t\text{-}1} - \beta(u_t - u^*_{t\text{-}1}) - \delta z_t + \eta_{1t}$$

où $u^*_{t\text{-}1}$ est le Nairu, $u_t$ est le taux de chômage, $\pi_t$ est l'inflation, $\Delta\pi_t$ est le taux de croissance de l'inflation, $\alpha(L)$ est un polynôme retard, $\delta$ est un vecteur de paramètres et $z_t$ est un vecteur de chocs d'offre temporaires. Ce modèle est dit triangulaire car l'inflation y trouve trois déterminants : l'inflation passée $(\alpha(L)\ \Delta\pi_{t\text{-}1})$, les tensions sur le marché du travail $(u_t - u^*_{t\text{-}1})$ et les chocs d'offre $(z_t)$. Les trois filtres ont pour objectif de fournir un Nairu plus variable avec le temps, mais ne permettent généralement pas d'expliquer le sens économique des évolutions du Nairu.

À l'origine, le filtre HP et le filtre HPMV estiment le chômage d'équilibre comme solution des programmes de minimisation suivants :

$$\underset{u^*_t}{Min} \sum_{t=1}^{T} \left(u_t - u^*_t\right)^2 + \lambda_1 \left(\Delta\Delta u^*_t\right)^2 \qquad \text{(HP)}$$

$$\underset{u^*_t}{Min} \sum_{t=1}^{T} \left(u_t - u^*_t\right)^2 + \lambda_1 \left(\Delta\Delta u^*_t\right)^2 + \lambda_2 \left(\eta_{1t}\right)^2 \qquad \text{(HPMV)}$$

18. Cette partie s'inspire très largement des travaux de Boone (2000) et Boone *et al.* (2001).

19. On pourra notamment se référer à Cogley et Nason (1995), à Boone (2000) ou à Guay et St Amant (1996).

20. Les filtres HPMV permettent d'intégrer des relations économiques au filtre HP (Laxton et Tetlow, 1992).

21. Même si seule la troisième méthode d'estimation est à proprement parler un filtre de Kalman, les filtres HP et HPMV peuvent aussi s'estimer à partir d'un tel filtre avec des contraintes sur les variances des innovations. Voir l'article « écart de production dans la zone euro : une estimation par le filtre HPMV ». de Chagny, Lemoine (2003) dans ce même dossier.

Ces filtres peuvent cependant se réécrire sous la forme de modèles espace-état. Cette réécriture permet d'utiliser une meilleure méthode d'estimation et de comparer la structure des différents modèles (tableau 4). En effet, ces trois méthodes peuvent être différenciées en fonction des équations de mesure et d'état spécifiées ainsi que des restrictions imposées sur les paramètres.

4. Résumé des différents modèles

| Modèles | Equations de mesure | Equations d'état | $\sigma^2_{\eta_1}$ | $\sigma^2_{\eta_2}$ | $\sigma^2_{\varepsilon}$ |
|---|---|---|---|---|---|
| HP | $u_t = u_t^* + \eta_{2,t}$ | $\Delta u_t^* = \Delta u_{t-1}^* + \varepsilon_t$ | - | 1/16 | 1/400 |
| HPMV | $\Delta\pi_t = \alpha(L)\Delta\pi_{t-1} - \beta(u_t - u_{t-1}^*) + \delta z_t + \eta_{1,t}$<br>$u_t = u_t^* + \eta_{2,t}$ | $\Delta u_t^* = \Delta u_{t-1}^* + \varepsilon_t$ | 1 | 1/16 | 1/400 |
| Kalman | $\Delta\pi_t = \alpha(L)\Delta\pi_{t-1} - \beta(u_t - u_{t-1}^*) + \delta z_t + \eta_{1,t}$ | $u_t^* = u_{t-1}^* + \varepsilon_t$ | R | - | Q |

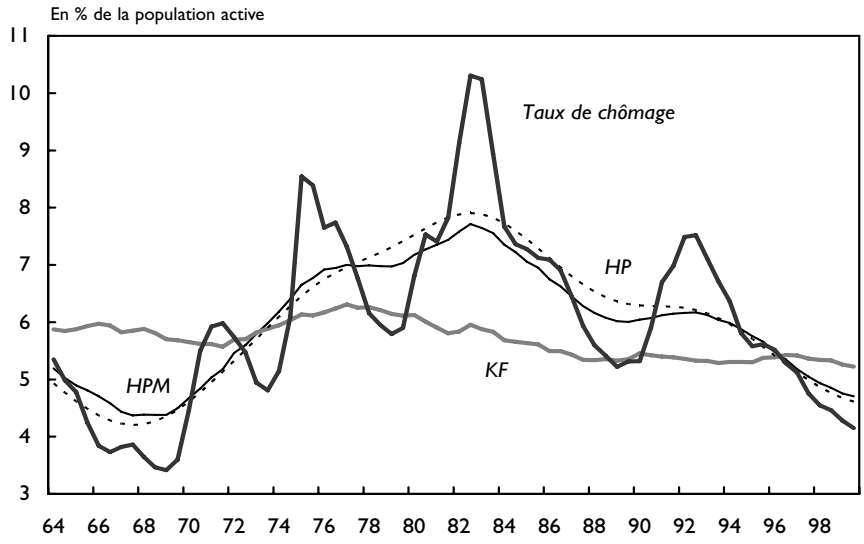
*Note* : Pour le filtre de Kalman, les variances des innovations de l'équation de mesure ($\sigma^2_{\eta 1}$) et de l'équation d'état ($\sigma^2_{\varepsilon}$) sont estimées librement avec les paramètres R et Q (selon les notations de l'encadré 1).

Pour le filtre HP, l'équation de mesure décrit le taux de chômage comme la somme d'une composante tendancielle (le taux de chômage d'équilibre) et d'un bruit, l'équation d'état définissant le taux de chômage d'équilibre comme une marche aléatoire en premières différences. Les variances des innovations et des erreurs de mesure sont fixées et égales à 1/400 et 1/16 respectivement, correspondant aux valeurs retenues pour une estimation avec des données semestrielles.

Pour le filtre HPMV, les résidus de la courbe de Phillips augmentée sont inclus dans le programme de minimisation. On utilise deux équations de mesure : la courbe de Phillips et le Nairu. L'équation d'état définit à nouveau la première différence du Nairu comme une marche aléatoire sans dérive. Les valeurs des variances de l'erreur de mesure, de l'innovation de la courbe de Phillips et des innovations du Nairu sont respectivement égales à 1, 1/16 et 1/400.

Pour le filtre de Kalman standard, la courbe de Phillips est directement utilisée comme équation de mesure. Il est à noter qu'aucune restriction n'est imposée sur la valeur des variances contrairement aux deux précédents modèles. Mais Stock et Watson (1996) ont montré que l'estimation par le filtre de Kalman pouvait biaiser vers 0 la variance des variables cachées d'une régression, soit ici la variance du Nairu $U^*$ qui apparaît dans la régression de l'inflation. Ce problème, dénommé le « *pile-up problem* », est illustré sur le graphique 4, le Nairu estimé par le filtre de Kalman ayant une variance excessivement faible.

4. Comparaison des filtres HP, HPMV et de Kalman




*Sources* : Bureau of Economic Analysis ; Bureau of Labor Statistics ; calculs des auteurs.

Les modèles espace-état sont appliqués à l'estimation du Nairu aux États-Unis sur la période 1963:1-1999:2 (graphique 4). Les estimations du Nairu évolutif par les filtres HP et HPMV fournissent des résultats convergents avec une baisse du chômage d'équilibre de presque 4 points entre le milieu des années 1980 à la fin des années 1990 où il passe sous les 5 %. En revanche, l'estimation du modèle « filtre de Kalman » est beaucoup moins variable, passant d'environ 6 % dans les années 1980 à un peu plus de 5 % dans les années 1990. Avec un chômage à environ 4 % en fin de période, cela signifierait que le chômage serait en 1999 à plus de 1 % sous le Nairu et impliquerait une forte accélération de l'inflation qui n'a pas été observée. L'estimation libre par le filtre de Kalman ne semble donc pas satisfaisante relativement aux estimations HP et HPMV. Ces mauvais résultats du filtre de Kalman sont liés au « *pile-up problem* », qui explique qu'on ait ici à tort un Nairu presque constant.

En conclusion, cet exemple a illustré une utilisation intéressante du filtre de Kalman, l'estimation de filtres HP ou HPMV, mis sous forme de modèles espace-état, avec un ratio contraint sur les variances. Mais cet exemple a également montré qu'un modèle espace-état ne peut pas toujours être estimé tel quel : la variance de la variable d'état doit parfois être calibrée sous peine d'obtenir des résultats biaisés avec une variable d'état quasi-constante.

# Contenu informatif de la courbe des taux sur l'inflation future [22]

Une dernière application concerne l'étude du contenu informatif de la structure par terme des taux d'intérêt pour l'inflation future. Ce type d'étude se justifie par l'intérêt des décideurs de la politique monétaire pour la courbe des taux comme une des variables financières permettant de prédire l'inflation. Mais la question n'est abordée ici qu'en tant qu'application possible des techniques espace-état. L'objectif est essentiellement de montrer, dans ce contexte, que le filtre de Kalman permet d'estimer un modèle linéaire à coefficients aléatoires.

Une méthode simple pour tester le contenu informatif (ou le pouvoir prédictif) de la courbe des taux pour l'inflation future est d'estimer l'équation suivante [23] par moindres carrés ordinaires (en utilisant éventuellement une correction de White ou de Newey et West) :

$$\pi_{t,m} - \pi_{t,n} = \alpha + \beta(i_{t,m} - i_{t,n}) + \eta_t \qquad (1)$$

où $\pi_{t,m}$ est le taux d'inflation à m-période mesuré par $log(P_{t+12m}/P_t)/m$, $P_t$ étant l'indice des prix à la consommation au mois $t$ et $i_{t,m}$ étant le taux d'intérêt nominal à m-période. Considérons ce modèle sur la période 1961:06-1996:12 pour le premier segment de la courbe des taux américains, soit le *spread* 2-1, ou encore l'écart entre le taux d'intérêt à deux mois et le taux à un mois. On obtient alors une équation dont le $R^2$ est assez faible avec une valeur de 0,20. La constante et le *spread* 2-1 ont des coefficients significatifs (tableau 5). Le rejet par un test de Student de l'hypothèse nulle $\beta = 0$ montre que ce segment de la courbe des taux possède un contenu informatif pour l'inflation future. De plus, l'intérêt a également porté sur l'égalité à 1 du paramètre $\beta$ pour tester la relation de Fisher [24]. Cette relation stipule que le taux d'intérêt réel *ex ante* entre $t$ et $t+m$ (égal au taux d'intérêt nominal à m-période moins l'inflation anticipée) est constant au cours du temps. Sur cette partie de la courbe des taux, l'hypothèse nulle $\beta = 1$ ne peut pas être rejetée et la relation de Fisher est donc validée.

5. Estimation par MCO

| Variables | Coefficients | Ecarts-types | t-stat |
|---|---|---|---|
| $\alpha$ | – 0,256 | 0,108 | – 2,369 |
| $\beta$ | 1,260 | 0,255 | 4,941 |

*Sources* : Federal Reserve Bank, OECD, Analytical Data Base ; calculs des auteurs.

22. Pour une présentation plus détaillée, voir Garbade et Wachtel (1978).
23. Pour une présentation détaillée, voir Mishkin (1990), Tzavalis et Wickens (1996).
24. Voir Kozicki et Tinsley (2001).

Cependant, l'inconvénient de cette méthode est qu'elle prédit en moyenne sur l'ensemble de la période un contenu informatif de la courbe des taux pour l'inflation future. Or, les travaux empiriques montrent que le contenu informatif des écarts de taux varie selon la période d'estimation retenue, le pays retenu et le segment considère de la courbe des taux. En particulier, cette relation a notamment été étudiée avec des tests de stabilité. Même si la date de rupture peut être endogénéisée, il est difficile de mettre en évidence plusieurs dates de rupture compte tenu des propriétés à distance finie de ces statistiques. Une approche alternative consiste à estimer l'équation précédente en faisant varier $\beta$. Par exemple, l'estimation de modèles à changements de régime permet d'envisager que ce coefficient prenne une valeur fixe sur certaines périodes et que cette valeur bascule avec les changements de régime (Chopin et Pelgrin, 2003). Il est alors possible de mettre en évidence plusieurs régimes. En d'autres termes, le contenu informatif pour l'inflation future varie dans le temps. Ces résultats sont confirmés par d'autres modèles non linéaires (modèles à seuils, non paramétriques). Finalement, les tests d'Andrews de fin d'échantillon montrent que le contenu informatif est faible pour les États-Unis lors de la dernière décennie.

Dans cette perspective, pour mettre en évidence la variabilité du contenu informatif de la courbe des taux pour l'inflation future, il convient d'utiliser une représentation espace-état simple où le coefficient représentant la pente de la relation (1) suit un processus stationnaire autour d'une moyenne [25]. Une autre alternative simple consisterait a implémenter la méthode des moindres carrés ordinaires récursive par une écriture espace-état.

À la suite de Garbade et Wachtel (1978), on envisage ici que ce coefficient puisse évoluer progressivement à l'aide d'un modèle linéaire à coefficient variable de la forme :

$$\begin{cases} \beta_{t+1} = \mu + \varphi_1 \beta_t + \varepsilon_t \\ \pi_{t,m} - \pi_{t,n} = \alpha + \beta_t (i_{t,m} - i_{t,n}) + \eta_t \end{cases} \quad (2)$$

avec une pente variant comme un simple processus auto-régressif d'ordre 1. En prenant $\beta_t$ comme variable d'état et $\pi_{t,m} - \pi_{t,n}$ comme variable de mesure, le modèle précédent est directement sous une forme espace-état.

Comme précédemment, les estimations sont menées sur le premier segment de la courbe des taux américaine sur la période 1961:06-1996:12. Pour estimer les paramètres du modèle espace-état (2), on applique l'algorithme EM et on obtient les valeurs présentées dans le tableau 6. Sur la base de cette estimation, on peut déterminer la valeur de long terme de la variable d'état $(\beta)$, soit $E[\beta]$=1,50. Par ailleurs,

25. Harvey (1989) utilise les modèles espace-état pour mettre en évidence et détecter l'instabilité dans une relation.

l'évolution de la variable d'état est représentée sur le graphique 5 et on constate que le contenu informatif de la courbe des taux pour l'inflation future varie au cours du temps. Les résultats sont donc cohérents avec la littérature existante. Cependant, il est a noter que l'incertitude entourant les estimations est assez importante. Cette variabilité peut être réduite en utilisant des méthodes statistiques plus sophistiquées.

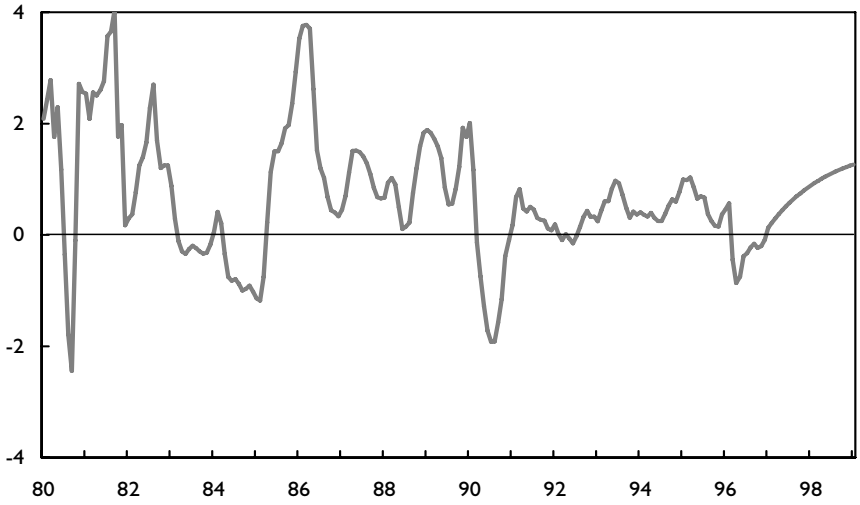
L'exercice simple précédent a montré comment le filtre de Kalman peut être utilisé pour estimer un modèle linéaire à coefficient aléatoire lorsque la relation étudiée peut être instable ou variable dans le temps. Il s'agirait sans doute d'approfondir la ou les sources d'instabilité (différentes représentations espace-état sont possibles avec des hypothèses différentes sur la constante ou la pente) et de tester par la suite le contenu informatif de la courbe des taux, ainsi que la relation de Fisher sur les sous-périodes considérées. Cependant, ce n'est pas l'objet de cet exemple.

6. Estimation du modèle espace-état

| Variables | Coefficients | Ecarts-types |
|---|---|---|
| $\alpha$ | – 0,14 | 0,03 |
| $\sigma_\varepsilon^2$ | 0,29 | 0,06 |
| $\mu$ | 0,09 | 0,07 |
| $\phi$ | 0,94 | 0,00 |
| $\sigma_\eta^2$ | 0,56 | 0,07 |

*Sources* : Federal Reserve Bank et OECD Analytical Data Base (ADBM) ; calculs des auteurs.

5. Évolution de $\beta_t$ sur la période 1980:1-1996:12




*Sources* : Federal Reserve Bank et OECD Analytical Data Base (ADBM) ; calculs des auteurs.

## Conclusion

Dans cet article ont été présentés les modèles espace-état et les techniques d'estimation associées, cadre de modélisation des séries temporelles issu des sciences de l'ingénieur. Ces modèles sont composés d'équations linéaires classiques (les équations de mesure) et supplémentaires servant à décrire la dynamique de variables cachées (les équations d'état). La procédure du maximum de vraisemblance consiste alors à estimer les variables cachées avec le filtre de Kalman et l'ensemble des paramètres du modèle avec l'algorithme EM.

Étant donné la généralité de la formalisation espace-état, de nombreux et divers modèles ou méthodes économétriques peuvent être reformulés et estimés par le maximum de vraisemblance dans ce cadre probabiliste : les modèles à composantes inobservables, les modèles à facteurs, les filtres HP ou HPMV et les modèles à coefficients aléatoires. Dans le cas des décompositions tendance-cycle, au lieu de recourir à une tendance coudée ou à un lissage par un filtre HP, le modèle à composante inobservable de Harvey (1989) détermine une tendance douce et un cycle, dont l'interprétation est facilitée par des paramètres de durée moyenne et de lissage. Pour estimer un indice coïncident d'activité, l'estimation d'un modèle à facteurs semble préférable à une moyenne pondérée arbitraire des composantes de l'activité. Pour éviter de recourir aux modèles structurels du Nairu dont les spécifications sont sujettes à de vifs débats, les techniques espace-état permettent également d'estimer le Nairu selon un filtre HPMV. Enfin, dans le cas d'équations instables au cours du temps, l'estimation d'un modèle à coefficients aléatoires permet de prendre en compte la variabilité des coefficients, par exemple du contenu informatif de la courbe des taux sur l'inflation future.

Mais ces applications ont également montré quelques difficultés qu'il convient de rappeler. Certains paramètres s'avèrent difficiles à identifier : la corrélation entre la tendance et le cycle a ici été supposée nulle, excluant tout effet d'hystérèse ; l'identification du taux de croissance moyen de l'indicateur coïncident a nécessité le développement d'une procédure spécifique. D'autre part, l'application au Nairu a montré la nécessité de calibrer la variance de la variable d'état qui risquerait d'être biaisée vers zéro. Enfin, si un modèle à coefficient aléatoire permet d'étudier le contenu informatif de la courbe des taux au cours du temps, il s'agirait de mieux prendre en compte l'incertitude qui entoure ces estimations.

## Références bibliographiques


BENTOGLIO G., J. FAYOLLE et M. LEMOINE, 2001 : « Unité et pluralité du cycle européen », *Revue de l'OFCE*, n° 78, pp. 9-73, juillet.

BOONE L., 2000 : *Comparing Semi-structural Methods to Estimate Unobserved Variables: The HPMV and Kalman Filters Approaches*, Economics Department Working Paper, OECD, n° 240.

BOONE L., M. JULLIARD, D. LAXTON et P. N'DIAYE, 2001 : *How Well Do Alternative Time Varying Parameter Models of the Nairu Help Policymakers Forecast Unemployment and Inflation in the OECD Countries?*, IMF Working Paper, février.

BOX G.E.P. et G.M. JENKINS, 1970 : *Times Series Analysis: Forecasting and Control*, San Francisco, CA: Holden-Day.

CHOPIN N. et F. PELGRIN, 2003 : « Bayesian Inference and State Number Determination for Hidden Markov Models: Application to the Information Content of the Yield Curve about Inflation », *Journal of Econometrics*, à paraître.

COGLEY T. et J.M. NASON, 1995 : « Effects of the HP filter on trend and difference stationary time series. Implications for business cycle research », *Journal of Economic Dynamics and Control*, n° 19, pp. 253-278.

DEMPSTER A.P., N.M. LAIRD et D.B. RUBIN, 1977 : « Maximum Likelihood for incomplete data via the EM algorithm », *Journal of the Royal Statistical Society*, Series B, vol. 39, pp. 1-38.

DIEBOLD F.X. et G.D. RUDEBUSH, 1999 : *Business Cycles, Durations, Dynamics and Forecasting*, Princeton University Press.

DURBIN J. et S.J. KOOPMAN, 2001 : *Time Series Analysis by State Space Methods*, Oxford University Press.

GARBADE K. and P. WACHTEL, 1978 : « Time Variation in the Relationship between Inflation and Interest Rates », *Journal of Monetary Economics*, Vol. 4.

GORDON R.J., 1997 : « The Time Varying Nairu and its Implications for Economic Policy », *Journal of Economic Perspectives*, Vol. 11, n° 1.

GOURRIEROUX C. et A. MONTFORT, 1990 : *Séries temporelles et modèles dynamiques*, Economica.

GUAY A. et P. ST-AMANT, 1996 : « Do Mechanical Filters Provide a Good Approximation of Business Cycle », *Technical Report* No. 78, Bank of Canada.

HARVEY A. C., 1989 : *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.

KALMAN R. E., 1960 : « A New Approach to Linear Filtering and Prediction Problems », *Journal of Basic Engineering, Transactions of the ASME*, Vol. 82, Series D.

KIM C. J. et C. R. NELSON, 1999 : *State-Space Models with Regime-Switching: Classical and Gibbs-Sampling Approaches with Applications*, MIT Press.

KOOPMAN S.J., N. SHEPARD et J.A. DOORNIK, 1998 : « Statistical Algorithms for Models in State Space Using SsfPack 2.2 », *Econometrics Journal*, vol. 1, pp. 1-55.

KOZICKI S. et P. TINSLEY, 2001 : « Term Structure Views of Monetary Policy Under Alternative Models of Agent Expectations », *Journal of Economic Dynamics and Control*, vol. 25, pp. 149-184.

LAXTON D. et R. TETLOW, 1992 : « A Simple Multivariate Filter for the Measurement of Potential Output », *Technical Report*, Bank of Canada, n° 59.

MISHKIN F., 1990 : « What Does the Term Structure Tell Us about Inflation? », *Journal of Monetary Economics*, 25, pp. 77-95.

MONTFORT A., 1997 : *Notes de cours sur les modèles dynamiques à facteurs*, ENSAE.

STOCK, J. H. et M. W. WATSON, 1991 : « A Probability Model of the Coincident Economic Indicators » in K. Lahiri and G. H. Moore, eds., *Leading Economic Indicators: New Approaches and Forecasting Records*. Cambridge University Press, pp. 63-89.

TZAVALIS E. et M.R. WICKENS, 1996 : « Forecasting Inflation from the Term Structure », *Journal of Empirical Finance*, 3, pp. 103-122.