



**Observatoire Economique et Statistique d'Afrique Subsaharienne**

**GUIDE METHODOLOGIQUE REVISE DE PRODUCTION DES  
COMPTES NATIONAUX TRIMESTRIELS DANS LES ETATS  
MEMBRES D'AFRISTAT**

**Juillet 2017**

## SOMMAIRE

<b><i>I. Introduction</i></b> .....	<b>3</b>
<b><i>II. Rappels du contenu du Guide méthodologique de 2011</i></b> .....	<b>3</b>
<b><i>III. Pourquoi réviser la méthodologie ?</i></b> .....	<b>4</b>
<b><i>IV. Révision méthodologique</i></b> .....	<b>4</b>
<b><i>V. Opérationnalisation</i></b> .....	<b>8</b>
<b><i>VI. Applications empiriques : exemples du Mali et du Bénin</i></b> .....	<b>9</b>
<b><i>Références bibliographiques</i></b> .....	<b>23</b>
<b><i>Annexes</i></b> .....	<b>24</b>

## I. Introduction

En 2011, l'Observatoire économique et statistique d'Afrique subsaharienne (AFRISTAT), avec l'appui de ses partenaires techniques et financiers, a mis à la disposition des Instituts nationaux de la statistique (INS) des Etats membres, une méthodologie d'élaboration des comptes nationaux trimestriels (CNT) selon l'approche économétrique basée sur la technique d'étalonnage-calage et l'approche mathématique. La technique de l'étalonnage-calage, la plus utilisée dans les Etats membres, repose sur un modèle de cointégration sur lequel est basée l'estimation des valeurs ajoutées trimestrielles.

Le présent guide révisé propose une amélioration de l'approche économétrique et est articulé autour de cinq sections, en dehors de l'introduction. La section 2 rappelle l'approche économétrique retenue dans le guide méthodologique de 2011. La section 3 justifie la révision du guide en mettant l'accent sur les raisons empiriques et statistiques. La section 4 présente la méthodologie révisée et ses implications en termes de prévisions. La section 5 a trait à l'opérationnalisation de la nouvelle méthode en 2 étapes relatives respectivement à l'estimation et à la prévision. La section 6 présente les résultats des applications empiriques sur les données du Mali et du Bénin ainsi que les leçons tirées de la mise en œuvre de ladite méthodologie.

## II. Rappels du contenu du Guide méthodologique de 2011

La technique de l'étalonnage-calage utilisée dans les Etats membres pour l'élaboration des comptes nationaux trimestriels consiste en deux étapes. La première, dite de l'étalonnage, vise à postuler d'abord une relation stable entre l'agrégat annuel et l'indicateur conjoncturel annualisé et à estimer les paramètres ; ensuite, à utiliser cette relation pour l'estimation des agrégats infra-annuels non observés. La deuxième étape, dite celle de calage, consiste à mettre en cohérence le compte annuel et le compte trimestriel de telle sorte à établir l'égalité entre la somme des agrégats infra-annuels estimés pour les quatre trimestres et le niveau de l'agrégat annuel. Pour ce faire, une redistribution du résidu de l'équation de la première étape est utilisée.

**1ère étape : estimation du modèle de base reposant sur l'équation suivante :**

$$y_n = \alpha + \beta Q_n + u_n, \text{ avec } n = 1, \dots, N$$

$y_n$  étant l'agrégat annuel et  $Q_n$  l'indicateur annualisé

$$Q_n = \sum_{t=1}^{4n} X_t, \text{ avec } t = 1, \dots, 4N$$

$X_t$  étant l'indicateur infra-annuel

L'indicateur annualisé étant supposé stationnaire, l'agrégat infra-annuel  $y_t$ ,  $Y_t$  est estimé en appliquant les méthodes des moindres carrés ordinaires (MCO) ou des moindres carrés généralisés (MCG) selon que les résidus sont supposés bruits blancs ou autocorrélés.

**2ème étape : estimation de l'agrégat infra-annuel et calage**

Une estimation de l'agrégat infra-annuel étalonné est obtenue par :

$$\hat{y}_t = \frac{\hat{\alpha}}{4} + \hat{\beta} Q_t, \text{ avec } t = 1, \dots, 4N$$

Pour assurer la cohérence entre l'agrégat annuel et la somme des estimations des agrégats trimestriels, les variations des résidus trimestriels sont lissées sous contrainte de cohérence entre le résidu annuel et la somme des résidus des quatre trimestres. Le résidu trimestriel est alors obtenu par lissage fondé sur le programme suivant :

$$\text{Min} \sum_{n=1}^N \sum_{t=4n-3}^{4n} (\hat{v}_t - \hat{v}_{t-1})^2$$

Sous contrainte 
$$\sum_{t=4n-3}^{4n} \hat{v}_t = \hat{u}_n \quad n = 1, \dots, N$$

Une estimation de l'agrégat infra-annuel étalonné et calé est obtenue par :

$$\hat{y}_t = \frac{\hat{\alpha}}{4} + \hat{\beta}Q_t + \hat{v}_t, \text{ avec } t=1, \dots, 4N$$

Dans la pratique, les Etats utilisent l'application informatique Ecotrim (conçue par EUROSTAT) qui fournit directement l'agrégat infra-annuel étalonné et calé.

### III. Pourquoi réviser la méthodologie ?

La révision méthodologique proposée part des insuffisances constatées dans le guide de 2011. En effet, en 2011, la méthodologie d'élaboration des comptes nationaux trimestriels repose sur l'approche économétrique basée sur la technique d'étalonnage-calage et l'approche mathématique. La technique de l'étalonnage-calage, la plus utilisée dans les Etats membres, repose sur un modèle de cointégration sur lequel est basée l'estimation des valeurs ajoutées trimestrielles.

Cette technique d'étalonnage-calage utilisée pour l'estimation de l'agrégat infra-annuel soulève quelques limites, pour plusieurs raisons :

- Sur le plan empirique, le choix des indicateurs infra-annuels pour trimestrialiser les comptes annuels montre un champ plus large des déterminants des valeurs ajoutées sectorielles à la lumière des conclusions du séminaire sur les comptes nationaux trimestriels organisé en 2015 à Bamako au Mali. Le tableau en annexe présente par branche d'activité le choix de ces indicateurs infra-annuels, à la lumière des pratiques observées dans les Etats membres qui élaborent les comptes nationaux trimestriels, ainsi que des recommandations qui en résultaient ;
- Sur le plan statistique, en relation avec la qualité des estimateurs, il est important de noter que :
  - o L'approche économétrique n'intègre pas le comportement de l'indicateur infra-annuel qui est en général non stationnaire. Cette non-stationnarité qui est conservée avec l'annualisation des données a des implications sur l'inférence des paramètres du modèle (en particulier sur le comportement asymptotique des estimateurs  $\hat{\beta}$ ) ;
  - o Le modèle utilisé dans la technique d'étalonnage-calage repose sur l'hypothèse que l'indicateur annualisé est exogène et stationnaire. Pourtant, on sait que la plupart des agrégats macroéconomiques sont non stationnaires (Nelson et Plosser, 1982). Par ailleurs, il est fort probable que l'indicateur annualisé soit endogène;
  - o Le calage qui implique une erreur d'estimation nulle sur l'indicateur conjoncturel annualisé couplé à la non prise en compte de la dynamique de l'indicateur infra-annuel rend difficile l'utilisation du modèle à des fins de prévision. Les tentatives de prévision qui ont été abordées avec ce modèle (BCEAO, 2014 ; INS Cameroun, 2012) sont basées sur des acquis de croissance qui supposent que le comportement de l'économie est inchangé à court-terme. Cette hypothèse paraît trop optimiste en situation de fortes fluctuations et occulte l'information disponible à la fin de chaque trimestre sur les indicateurs conjoncturels.

### IV. Révision méthodologique

La démarche proposée constitue une amélioration de la méthodologie développée par AFRISTAT en ce qu'elle intègre l'endogénéité des indicateurs infra-annuels de conjoncture considérés comme principaux déterminants des agrégats annuels que sont les valeurs ajoutées sectorielles. Elle se distingue donc de la technique d'étalonnage-calage par le fait qu'elle intègre dans l'équation de la

mesure de l'agrégat annuel, le niveau et la variation de l'annualisé de l'indicateur infra-annuel à laquelle est associée une deuxième équation dynamique de l'indicateur infra-annuel.

Cette démarche repose sur deux étapes : une première étape qui vise à postuler une relation linéaire entre les variables observées à rythme annuel et d'en estimer les paramètres et une seconde étape qui utilise lesdits paramètres pour proposer des estimations et prévisions à court-terme de variable dépendante inobservable en exploitant la dynamique des indicateurs infra-annuels utilisés.

#### IV.1. Modèle et estimation des paramètres

Soient  $\{y_t\}$  un processus univarié inobservable à rythme trimestriel mais observable à rythme annuel et  $\{X_t\}$  un processus observable à n'importe quelle fréquence. On désigne par  $\{y_n\}$  et  $\{Q_n^S\}$  respectivement les processus observables à rythme annuel de  $\{y_t\}$  et  $\{X_t\}$  tel que  $Q_n^S = \sum_{i=1}^S X_{S(n-1)+i}$  avec  $S = 4$ .

Ils sont générés par le modèle défini par les équations (1) et (2) suivantes :

$$y_n = \alpha_0 + \beta_0 Q_n^S + u_n, \quad n = 1, 2, \dots, N \quad (1)$$

et

$$X_t = \exp(cS/T) X_{t-S} + \varepsilon_t, \quad t = 1, 2, \dots, SN \quad (2)$$

avec  $N$  le nombre d'années,  $S = 4$  le nombre de trimestres dans une année,  $T = SN$  le nombre total d'observations trimestrielles et  $\varepsilon_t$  l'erreur aléatoire supposée stationnaire de densité spectrale à la fréquence 0 vérifiant  $\sigma_{\varepsilon_t}^2 = 2\pi f_{\varepsilon_t}(0)$ .

On peut donc écrire  $X_t$  en introduisant, comme dans HEGY (1990), la période  $s$  et l'année  $n$  comme ci-après  $X_t = X_{sn} = X_{S(n-1)+s}$  avec  $s = 1, 2, \dots, S$  et  $n = 1, 2, \dots, N$ . Le processus  $\{X_t\}$  est saisonnier et quasi-intégré. Avec l'approximation  $\exp(c/N) \approx 1 + c/N$ , le paramètre  $c$  s'interprète comme la mesure de la déviation du processus  $X_t$  par rapport à la racine unitaire saisonnière. En effet, si  $c = 0$ ,  $\{X_t\}$  est un processus autorégressif saisonnier d'ordre 1,  $SAR(1)$ .

Si  $c < 0$  alors  $\{X_t\}$  est un processus localement stationnaire. Il est saisonnier et localement explosif si  $c > 0$ .

Dans le modèle (1), on suppose aussi que l'erreur aléatoire  $u_n$  est stationnaire et corrélée avec  $v_n = \sum_{i=1}^S \varepsilon_{S(n-1)+i}$ . La corrélation entre  $u_n$  et  $v_n$  implique celle entre  $Q_n^S$  et  $u_n$ . De ce fait,

$Q_n^S$  est endogène. On pose  $\omega_n = (u_n, v_n)'$  alors  $\omega_n$  est un vecteur de processus faiblement stationnaires  $I(0)$  vérifiant  $E(w_n) = 0$  pour tout  $n$  avec les hypothèses classiques d'existence des moments. Si  $c = 0$ ,  $\{Q_n^S\}$  est  $I(1)$ . Et si  $\beta_0 \neq 0$ , on en déduit sous l'hypothèse de stationnarité de  $\omega_n$  que le système  $\{y_n, Q_n^S\}'$  est cointégré avec  $(1, -\beta)$  comme vecteur de cointégration.

La densité spectrale  $f_{\omega_t}(0)$  à la fréquence 0 de  $\omega_t$  est donnée par la matrice  $\Omega$  définie par

$$\Omega = 2\pi f_{\omega_t}(0) = \sum_{j=-\infty}^{\infty} E\left(\omega_t \omega_{t-j}'\right) = \begin{bmatrix} \omega_{11} & \omega_{21} \\ \omega_{12} & \omega_{22} \end{bmatrix} > 0$$

avec  $\omega_{12} = \omega_{21}$ . Si  $u_n$  et  $v_n$  n'étaient pas corrélés, alors  $\omega_{12} = \omega_{21} = 0$  et  $Q_n^S$  serait exogène.

Considérons  $L$ , la matrice triangulaire supérieure définie par

$$L = \begin{bmatrix} 1 & -\omega_{12}\omega_{22}^{-1} \\ 0 & 1 \end{bmatrix} \quad (3)$$

et  $(\varepsilon_{u.v_n}, v_n) = L \omega_n$ . Le processus aléatoire  $\varepsilon_{u.v_n}$  est orthogonal à  $v_n$  au sens où la densité spectrale à la fréquence 0 est une matrice diagonale. Celle-ci vaut:

$$2\pi f_{\left(\varepsilon_{u.v_n}, v_n\right)}(0) = \begin{bmatrix} \omega_{11} - \omega_{12}^2 \omega_{22}^{-1} & 0 \\ 0 & \omega_{22} \end{bmatrix} > 0$$

On en déduit que le modèle (1) est équivalent à:

$$y_n = \alpha_0 + \beta_0 Q_n^S + \gamma_0 \Delta_c Q_n^S + \varepsilon_{u.v_n}, \quad n=1, 2, \dots, N \quad (4)$$

avec  $\Delta_c Q_n^S = Q_n^S - \exp(c/N) Q_{n-1}^S = v_n$  et  $\gamma_0 = \omega_{12} \omega_{22}^{-1}$ .

Pour l'estimation des paramètres, on considère la régression augmentée ci-après:

$$y_n = \alpha + \beta Q_n^S + \gamma \Delta_c Q_n^S + \eta_n \quad (5)$$

Sous l'hypothèse d'endogénéité de  $Q_n^S$ , la régression (5) présente deux avantages. En premier lieu, elle permet de faire l'inférence sur  $\beta$  en se basant sur l'approche standard fondée sur la normalité (voir annexe et aussi Phillips et Hansen (1990) et Jansson (2002) ); en second lieu, elle permet d'exploiter la dynamique de  $Q_n^S$  pour prédire les futurs de  $y_n$ .

Les estimateurs des moindres carrés ordinaires de  $\alpha$ ,  $\beta$  et  $\gamma$  sont donnés par  $\hat{\Gamma} = (X'X)^{-1} X'y$  avec  $\hat{\Gamma} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})'$ ,  $X$  est la matrice  $N \times 3$  des variables explicatives telle que  $X = [1 \ Q^S \ \Delta_c Q^S]$  et  $y$  le vecteur des observables  $y_n$ .

Sous l'hypothèse nulle  $H_0 : \beta = \beta_0 \neq 0, \gamma = \gamma_0$ , on a

$$\hat{\Gamma} - \Gamma_0 = (X'X)^{-1} X' \varepsilon_{uv} \quad (6)$$

avec  $\Gamma_0 = (\alpha_0, \beta_0, \gamma_0)'$  et  $\text{var}(\hat{\Gamma}) = \sigma_{\varepsilon_{uv}}^2 (X'X)^{-1}$

L'estimateur  $\hat{\alpha}$  des moindres carrés ordinaires du paramètre  $\alpha$  est convergent et les estimateurs  $\hat{\beta}$  et  $\hat{\gamma}$  de  $\beta$  et  $\gamma$  sont hyperconvergents au sens où d'une part,  $\sqrt{N}(\hat{\alpha} - \alpha_0)$  et  $N(\hat{\beta} - \beta_0)$

sont asymptotiquement gaussiens et  $N(\hat{\gamma} - \gamma_0)$  converge vers une loi non standard et, d'autre part,  $(\hat{\beta} - \beta_0) = o_P(1)$ ,  $(\hat{\alpha} - \alpha_0) = o_P(1)$ . et  $(\hat{\gamma} - \gamma_0) = o_P(1)$

(Voir proposition 2 en annexe).

#### IV.2 Estimation et Prévision infra-annuelles de l'inobservable $y_t$

L'estimation de la variable annualisée  $y_n$  fondée sur la régression (5) est donnée par

$$\hat{y}_n = \hat{\alpha} + \hat{\beta} Q_n^S + \hat{\gamma} \Delta_c Q_n^S \quad (7)$$

avec comme erreur d'estimation  $y_n - \hat{y}_n$  et l'erreur quadratique moyenne

$RMSE = \left[ N^{-1} \sum_{n=1}^N (y_n - \hat{y}_n)^2 \right]^{1/2}$  qui converge vers  $\sigma_{\varepsilon_{u,v_n}}$  lorsque  $N$  est de plus en plus grand.

Pour estimer la variable infra-annuelle inobservable  $y_t = y_{S(n-1)+s}$  sur la période d'observation des données annuelles, on fait l'hypothèse que la variable  $y_{S(n-1)+s}$  sous la contrainte

$y_n = \sum_{s=1}^S y_{S(n-1)+s}$  obéit de façon implicite au processus générateur de données similaire à (2).

On en déduit l'estimation  $\hat{y}_{S(n-1)+s}$  suivante pour  $y_{S(n-1)+s}$

$$\hat{y}_{S(n-1)+s} = \hat{\alpha} / S + \hat{\beta} X_{S(n-1)+s} + \hat{\gamma} \left[ X_{S(n-1)+s} - \exp(c/N) X_{S(n-2)+s} \right] \quad (8)$$

avec la contrainte  $\hat{y}_n = \sum_{s=1}^S \hat{y}_{S(n-1)+s}$

On s'intéresse à la prévision de  $y_t$  à la date  $t+h$  conditionnellement à l'information disponible à la date  $t$  en exploitant le modèle (2) avec la même hypothèse que ci-dessus. Ainsi, pour l'année  $n$ , on peut écrire  $y_t$  en fonction de  $n, S$  et de la saison  $s$  ( $s=1, \dots, S$ ) de la façon suivante :

$y_t = y_{S(n-1)+s}$ . L'information disponible à la date  $n$ , est constituée des passés de  $y_n$ , soit

$y_{n-1}, y_{n-2}$  et des données trimestrielles correspondantes  $X_{S(n-2)+s}, X_{S(n-3)+s}$  avec

$s=1, 2, \dots, S$  auxquelles s'ajoutent pour l'année courante les  $X_{S(n-1)+h}$  pour  $h=1, \dots, s$ .

Sur la base de l'expression (2) et de la proposition 1, on propose pour  $y_{S(n-1)+s}$  la prévision

suivante à la date  $S(n-1)+s+h$ :

$$\hat{y}_{\{S(n-1)+s+h\}} / I_t = \begin{cases} \hat{\alpha} / S + \hat{\beta} X_{\{S(n-1)+s\}} + \hat{\gamma} [X_{\{S(n-1)+s\}} - \exp(c/N) X_{\{S(n-2)+s\}}] & \text{si } h=0 \\ \hat{\alpha} / S + \hat{\beta} \exp(c/N) X_{\{S(n-2)+s+h\}} & \text{si } 1 \leq h \leq S-s \end{cases} \quad (9)$$

On en déduit que, pour tout  $s = 1, \dots, S$ , la prévision annuelle de  $y_n$  fondée sur l'information disponible à la période  $t = S(n-1) + s$  est donnée par  $\hat{y}_n / I_t = \Psi_s \hat{\Gamma}$  avec  $\Psi_s$  la matrice  $1 \times 3$  définie par  $\Psi_s = [1 \quad Z_{1,s} \quad Z_{2,s}]$

$$\text{où } Z_{1,s} / I_t = \begin{cases} \exp(c/N) Q_{n-1}^S + \sum_{h=1}^s \Delta_c X_{\{S(n-1)+h\}} & \text{si } s < S \\ Q_n^S & \text{si } s = S \end{cases} \quad (10)$$

$$\text{et } Z_{2,s} / I_t = \begin{cases} \sum_{h=1}^s \Delta_c X_{\{S(n-1)+h\}} & \text{si } s < S \\ \Delta_c Q_n^S & \text{si } s = S \end{cases} \quad (11)$$

L'erreur de prévision  $e_n = y_n - \hat{y}_n / I_t$  de la variable annuelle  $y_n$  fondée sur l'information disponible à la fin de la saison  $s$  de l'année  $n$  a comme variance  $\sigma_e^2$  donnée par

$$\sigma_e^2 = \text{var}(e_n) = \sigma_{\hat{\varepsilon}_{u.v}_n}^2 \left( 1 + \Psi_s (X'X)^{-1} \Psi_s' \right) \quad (12)$$

On en déduit, l'intervalle de confiance  $I(y_n)$  au seuil  $(1 - \alpha)$  ci-après:

$$I(y_n) = \left] \hat{y}_n / I_t - \sigma_e t_{1-\alpha/2}; \hat{y}_n / I_t + \sigma_e t_{1-\alpha/2} \right[ \quad (13)$$

où  $t_{1-\alpha/2}$  est le quantile d'ordre  $1 - \alpha/2$  de la loi normale centrée et réduite  $N(0,1)$ .

## V. Opérationnalisation

La mise en place de comptes nationaux trimestriels doit permettre de produire et publier des données fiables de façon régulière. Elle repose sur deux phases successives, une phase **préparatoire** et une phase de **production régulière**.

La phase préparatoire a pour objectif de tester la possibilité de produire des comptes nationaux trimestriels fiables et précis dans des délais répondant aux besoins des utilisateurs. Cette phase exige de disposer d'une base d'informations assez fournies sur les comptes annuels et les indicateurs infra-annuels pour assier le dispositif d'élaboration des comptes nationaux trimestriels.

La seconde phase est le passage à la production et à la diffusion régulière de comptes trimestriels sur l'année courante. La démarche révisée propose à ce niveau de faire les estimations et les prévisions des comptes nationaux trimestriels.

### V.1 Etape d'estimation

Dans le cadre de la méthodologie révisée pour l'estimation des comptes nationaux trimestriels, la démarche est la suivante :

- Construire des graphiques pour analyser l'évolution de l'indicateur infra-annuel et de l'agrégat annuel ;

- Procéder à des tests de racines unitaires saisonnières sur les indicateurs infra-annuels ;
- Annualiser les indicateurs infra-annuels, ainsi que les résidus issus des modèles de régression caractéristiques de l'indicateur infra-annuel ;
- Construire les graphiques concomitamment des agrégats annuels, des indicateurs annualisés, ainsi que des résidus annualisés pour détecter la cointégration ;
- Procéder à un test de cointégration à la Perron-Rodriguez 2016 (approche GLS-Perron).

#### **IV.2 Etape de prévision**

Les travaux à réaliser dans cette phase s'appuient sur les résultats des estimations des agrégats du dernier trimestre observé, afin de prédire les niveaux et les évolutions des trimestres en cours et à venir. La démarche consiste à :

- Prédire les indices des trimestres à venir sur la base de l'information disponible sur le trimestre qui vient de s'achever, afin de couvrir une année complète ;
- Prédire la valeur ajoutée des branches pour l'année, en fonction de l'information disponible sur le trimestre qui vient de s'achever ;
- En déduire le PIB annuel fondé sur l'information disponible à partir des résultats des estimations et de la prévision.

## **VI. Applications empiriques : exemples du Mali et du Bénin**

Les données utilisées sont celles provenant des instituts nationaux de statistique du Bénin sur la période 2000-2015 et du Mali sur celles plus réduites de 2003-2015. Elles portent sur les composantes sectorielles de l'indice harmonisé de production industrielle (IHPI) et du produit intérieur brut à prix constants (PIB).

En vue de réaliser les prévisions de valeurs ajoutées trimestrielles pour l'année 2015, les périodes d'estimation des paramètres du modèle et des valeurs ajoutées trimestrielles sont réduites à 2000-2014 pour le Bénin et 2003-2014 pour le Mali.

### **VI.1 Mise en évidence du comportement saisonnier des indicateurs infra-annuels**

Avant d'estimer les paramètres du modèle proposé dans cette étude, on analyse par le biais des tests de racine unitaire saisonnière, le comportement en termes de saisonnalité de quelques composantes sectorielles de l'indice harmonisé de la production industrielle (IHPI) fondés sur le modèle de régression de HEGY(1990) dont l'extension est proposée par Ghysels et al. (1994).

L'existence de racine unitaire à la fréquence 0 dans  $x_t$  est mise en évidence par le test de l'hypothèse nulle  $H_0 : \pi_1 = 0$  contre l'alternative de stationnarité  $H_1 : \pi_1 < 0$ . La présence de racine unitaire saisonnière à la fréquence  $\pi$  est mise en exergue par le test de l'hypothèse nulle  $H_0 : \pi_2 = 0$  contre  $H_1 : \pi_2 < 0$ . Celle de l'existence de racines unitaires saisonnières conjuguées ( $\pm i$ ) correspondant aux fréquences  $\pi/2$  et  $3\pi/2$  est testée par l'hypothèse nulle jointe  $H_0 : \pi_3 = \pi_4 = 0$  contre l'alternative  $H_1 : \pi_3 \neq 0$  ou  $\pi_4 \neq 0$ . Les valeurs critiques des statistiques de student et des statistiques de Fischer des hypothèses jointes sont tabulées par Dickey et al (1984), HEGY (1990) et Ghysels et al (1994).

Les tableaux 5a et 5b présentent les statistiques de Student pour les  $\pi_i$  ainsi que les statistiques de Fischer des tests conjoints. L'analyse des résultats montre que l'existence de racine unitaire à la fréquence 0 n'est pas rejetée au niveau conventionnel de 5% pour toutes les series. En outre, pour l'existence de racine unitaire saisonnière, le comportement des variables présente des similitudes et aussi des différences selon les pays. Le test de l'hypothèse conjointe d'existence de racines unitaires saisonnières aux fréquences  $\pi/2$  et  $3\pi/2$  n'est pas rejetée au niveau de 5% pour l'indice harmonisé

de production industrielle (IHPI) de la branche "industrie alimentaire" pour les deux pays. L'IHPI des industries chimiques au Bénin et L'IHPI des industries d'égrenage de coton au Mali, de même que l'IHPI global présentent des racines unitaires saisonnières à toutes les fréquences. La présence de racine unitaire saisonnière à la fréquence  $\pi$  n'est pas rejetée pour l'IHPI des industries "eau et électricité" au Mali. Ce résultat contraste avec celui du Bénin par le rejet de cette même hypothèse. Des similitudes sont relevées également pour les industries textiles avec le rejet de la présence des racines unitaires saisonnières à toutes les fréquences. Au total, les résultats suggèrent pour la plupart des séries, la présence de racine unitaire saisonnière.

**Tableau 5a:** Statistiques de Student et de Fischer pour les tests de racine unitaire saisonnière de quelques composantes sectorielles de l'indice de production industrielle (IHPI) du Bénin

modèle:  $\Delta_4 x_t = \mu + \gamma t + \sum_{i=1}^3 \delta_i SD_{it} + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \sum_{j=1}^l \Gamma_j \Delta_4 x_{t-j} + e_t$   
avec  $\Delta_4 = 1 - B^4$  où  $Bx_t = x_{t-1}$ ,  $y_{1,t} = (1+B+B^2+B^3)x_t$ ,  $y_{2,t} = -(1-B+B^2-B^3)x_t$ ,  $y_{3,t} = -(1-B^2)x_t$  et  $SD_{it} = 1$  si  $i = t \pmod{S}$  et  $SD_{it} = 0$  sinon.

Pays: Bénin; période 2000q1:2015q4	lag l	t[Pi1]	t[Pi2]	t[Pi3]	t[Pi4]	F[3-4]	F[2-4]	F[1-4]
1. IHPI Industries alimentaires	2	-3,055	-3,482	-2,463	-1,637	5,023	9,974	10,331
2. IHPI Industries textiles	0	-2,038	-3,937	-5,146	-1,145	14,568	16,124	13,085
3. IHPI Industries chimiques	2	-1,337	-3,023	-1,271	-1,553	2,022	5,038	4,082
4. IHPI Industries eau, gaz et électricité	1	-0,960	-3,818	-3,424	-2,670	12,008	20,050	15,167
5. IHPI Autres industries	1	-3,109	-5,611	-5,155	-1,179	15,050	25,997	22,103
6. IHPI global	2	-2,142	-3,152	-2,214	-1,643	4,453	8,579	7,852
Valeurs critiques								
5%	2	-3,675	-3,053	-3,625	-1,916	6,560	6,071	6,518
	1	-3,668	-3,048	-3,618	-1,917	6,562	6,067	6,516
	0	-3,050	-3,017	-3,571	-1,975	6,593	6,058	5,909
10%	2	-3,341	-2,711	-3,253	-1,486	5,399	5,130	5,704
	1	-3,335	-2,707	-3,248	-1,487	5,405	5,130	5,703
	0	-2,699	-2,676	-3,217	-1,530	5,514	5,165	5,004

Sources: Nos calculs, données INSAE, Bénin

**Tableau 5b:** Statistiques de Student et de Fischer pour les tests de racine unitaire saisonnière de quelques composantes sectorielles de l'indice de production industrielle (IHPI) du Mali

modèle:  $\Delta_4 x_t = \mu + \gamma t + \sum_{i=1}^3 \delta_i SD_{it} + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \sum_{j=1}^l \Gamma_j \Delta_4 x_{t-j} + e_t$   
avec  $\Delta_4 = 1 - B^4$  où  $Bx_t = x_{t-1}$ ,  $y_{1,t} = (1+B+B^2+B^3)x_t$ ,  $y_{2,t} = -(1-B+B^2-B^3)x_t$ ,  $y_{3,t} = -(1-B^2)x_t$  et  $SD_{it} = 1$  si  $i = t \pmod{S}$  et  $SD_{it} = 0$  sinon.

Pays: Mali ; période: 2003q1:2015q4	lag l	t[Pi1]	t[Pi2]	t[Pi3]	t[Pi4]	F[3-4]	F[2-4]	F[1-4]
1. IHPI industries d'égrenage de coton	1	-1,012	-2,197	-1,748	-1,161	2,434	3,470	2,813
2. Industries extractives	2	-3,540	-4,247	-4,447	-0,504	10,213	23,745	26,581
3. IHPI industries agro-alimentaires	2	-1,267	-1,233	-0,616	-0,427	0,249	1,022	1,349
4. IHPI industries métallurgie fonderie	2	-2,069	-3,297	-3,926	-2,495	14,034	18,341	15,506
5. Industries de textiles et d'habillement	1	-3,356	-3,735	-5,272	-0,997	15,628	17,454	15,675
6. IHPI Industries Eau et électricité	2	-1,828	-2,298	-1,380	-3,079	8,175	11,306	9,217
7. IHPI Autres industries	1	-3,342	-4,429	-5,100	-1,572	16,244	22,220	18,976
8. IHPI global	2	-1,631	-2,894	-2,213	-1,303	3,492	7,143	6,495
Valeurs critiques								
5%	1,2	-3,710	-3,080	-3,660	-1,910	6,550	6,090	6,530
10%	1,2	-3,370	-2,730	-3,280	-1,480	5,370	5,130	5,710

Sources: Nos calculs, données INSTAT, Mali

## **VI.2 Estimation de valeurs ajoutées trimestrielles et comparaison avec les valeurs annuelles observées**

L'estimation des paramètres se fait en deux étapes. La première estime le modèle d'état sur les données trimestrielles et en récupère les paramètres d'intérêt. La seconde porte sur le modèle dit « de mesure » sur les données annuelles. L'attention étant focalisée sur la significativité des paramètres du modèle, en particulier  $\alpha$ ,  $\beta$ , et  $\gamma$  pour certaines branches, on note que le paramètre de l'indice harmonisé de production industrielle (IHPI) de la branche n'est pas significatif, ce qui a amené à remplacer l'IHPI de la branche retenue par l'IHPI d'une autre branche ou l'IHPI global pour exploiter l'interdépendance entre celles-ci. Cependant, tenant compte de la corrélation entre les IHPI des branches, les variables « étalon » ou explicatives au sens de la significativité des paramètres d'intérêt, ont pu être identifiées.

La non significativité des coefficients de l'IHPI de la branche concernée pourrait s'expliquer d'une part, par la faible représentativité de l'échantillon ayant servi aux calculs de cet indicateur, et d'autre part, par la non prise en compte du secteur informel qui s'avère important dans certaines branches.

Les tests de racine unitaire sont faits sur les résidus ( $\hat{\varepsilon}_t$ ) de l'équation (2) en utilisant l'approche GLS-detrending de Elliot et al (1996) avec les statistiques de la classe MZ reconnues pour leur bon niveau et leur bonne puissance dans les petits échantillons. S'agissant des tests de cointégration qui reposent sur les tests de racine unitaire sur les résidus ( $\hat{\eta}_n$ ) de l'équation de mesure (5), ils ont été réalisés avec la statistique Z en exploitant les travaux récents de Perron et Rodriguez (2016) développés avec l'approche GLS detrending. Les résultats présentés en annexe, montrent qu'au niveau conventionnel de 5%, l'hypothèse nulle de non cointégration est rejetée au profit de celle de la cointégration.

Les tableaux 2 et 3 présentent respectivement les résultats pour le Bénin et le Mali. On note, en général, l'impact significatif à la fois de l'indicateur annualisé de l'IHPI et de sa variation sur la valeur ajoutée sectorielle relative à la branche retenue.

**Tableau 2:** Estimation des paramètres du modèle de régression sur données annuelles du Bénin

Branche	Variable "étalon"	Paramètres estimés			variance/densité spectrale			Nombre d'observations (N)	R <sup>2</sup>	
		$\alpha$	$\beta$	$\gamma$	exp(c/n)	$v_n$	$\varepsilon_{u.v_n}$			
Agriculture	ihpi global		177,89	0,73	-0,703	1,04	1,06	1,5686	14	0.893
		t-stat normalisé	7,12 ***	12,86 ***	-12,508 ***					
Transports, postes et télécommunications	ihpi global		-124,66	0,83	-0,083	1,04	1,06	0,7244	14	0.924
		t-stat normalisé	-2,67 **	7,84 ***	-0,801					
Electricité, gaz et eau	ihpi Electricité et eau		-17,77	0,08	0,024	1,08	1,12	0,51	14	0.934
		t-stat normalisé	-2,82 **	6,21 ***	2,112 **					
BTP	ihpi autres industries		87,59	0,28	-0,105	1,02	1,43	0,4816	14	0.557
		t-stat normalisé	1,31	1,72 *	-2,494 **					
PIB_Comres-s	ihpi autres industries		158,33	0,54	-0,424	1,02	1,43	0,6939	14	0.617
		t-stat normalisé	1,81 *	2,50 **	-7,743 ***					
Industries Agro-alimentaires	ihpi industries agro-alimentaires		417,28	-0,35	0,065	1,03	1,20	0,56	14	0.516
		t-stat normalisé	5,1408 ***	-1,9096 *	-0,882					

Source: Nos calculs, données INSAE, Bénin; Période 2001-2014, (\*\*),(\*\*),(\*)= significatifs à 1%,2% et 10%

**Tableau 3** : Estimation des paramètres du modèle de régression sur données annuelles du Mali

Branche	Variable "étalon"		Paramètres estimés				variance/densité spectrale			Nombre d'observations (N)	R <sup>2</sup>
			$\alpha$	$\beta$	$\beta$	$\gamma$	exp(c/n)	$v_n$	$\varepsilon_{u,v_n}$		
R2Agriculture d'exportation	IHPI Extraction d'or		186,42	-0,321	0,288		0,8897	1,38	1,4502	11	0.412
		t-stat normalisé	0,22	-25,494 ***	17,044 ***						
Activités extractives	IHPI global		66,92	-0,081	-0,094		0,9836	1,141	0,0627	11	0.404
		t-stat normalisé	0,04	-7,216 ***	-5,153 ***						
Industries agroalimentaires	Fabrication de produits alimentaires		264,58	-0,314	0,678		0,9777	0,7794	0,8385	11	0.510
		t-stat normalisé	0,06	-6,885 ***	32,773 ***						
Textiles	IHPI composite "Fabrication de textiles" et "Fabrication d'article d'habillement, teinture fourrure"		-125,30	0,562	-0,397		0,9458	1,417	0,2249	11	0.500
		t-stat normalisé	-0,01	8,008 ***	-15,902 ***						
Métallurgie fonderie (Or)	IHPI Extraction d'or		710,86	-1,472	0,734		0,8897	1,38	5,2378	11	0.787
		t-stat normalisé	0,36	-51,298 ***	36,183 ***						
Autres industries	IHPI Extraction d'or		117,46	-0,125	0,071		0,8897	1,38	0,2704	11	0.445
		t-stat normalisé	0,41	-30,172 ***	5,493 ***						
Electricité et eau	IHPI composite "production d'eau et électricité" et "Captage, traitement et distribution d'eau"		53,50	-0,028	-0,011		1,075	1,1076	0,0327	11	0.558
		t-stat normalisé	0,22	-80,268 ***	-1,164 **						
Construction	IHPI global		-13,04	0,108	-0,138		0,9777	1,141	2,5132	11	0.314
		t-stat normalisé	-0,01	11,382 ***	-49,319 ***						

Source: Nos calculs, données INSTAT, Mali; Période 2004-2014, (\*\*\*),(\*\*),(\*)= significatifs à 1%,2% et 10%

Ainsi, pour le Mali, l'augmentation de l'IHPI de la branche « extraction d'or », fait diminuer la production des branches « Agriculture d'exportation », « métallurgie fonderie (or) » et « autres industries ». En revanche, la variation de cet indice sur deux années consécutives accroît la production des branches considérées. L'ampleur de cette variation couplée avec le niveau annuel de l'indice est déterminante pour l'appréciation de la variation de la production desdites branches. Des résultats similaires sont obtenus pour l'ensemble des branches avec des paramètres alfa et gamma significatifs et de signes opposés à l'exception de la branche « activités extractives ».

Pour les données du Bénin, les paramètres d'intérêt estimés sont aussi significatifs et de signes opposés pour toutes les branches, sauf pour la branche « électricité, gaz et eau » où les signes sont tous positifs avec la variable explicative identique à l'IHPI de ladite branche.

L'estimation des valeurs ajoutées trimestrielles à l'aide des paramètres estimés et de l'information disponible sur les indicateurs infra-annuels a été réalisée sur la période d'observation. L'agrégation annuelle qui en résulte est comparée à la valeur ajoutée observée par branche. Les graphiques 1 et 2 présentent l'évolution simultanée des valeurs observées et estimées. La synthèse de la comparaison est donnée par l'erreur quadratique moyenne comparée à la moyenne de la valeur ajoutée observée (cf. tableau 4).

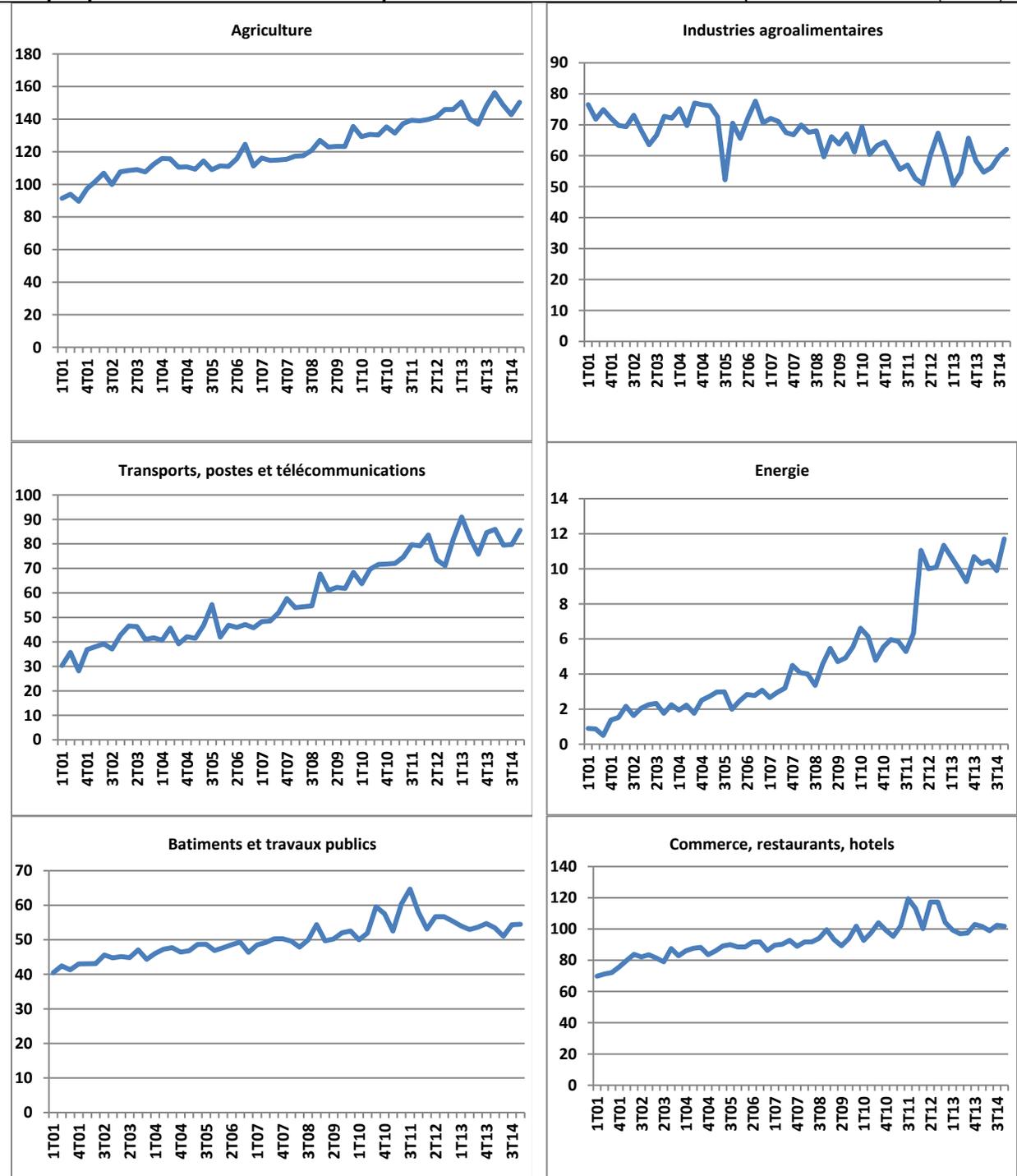
**Tableau 4** : Erreur quadratique moyenne rapportée à la moyenne de la valeur ajoutée observée

Mali		Bénin	
Branche	EQM / moyenne (VA)	Branche	EQM /moyenne(VA)
Agriculture d'exportation	9,87	Agriculture	1,01
Activités extractives	1,15	Industries agro-alimentaires	1,96
Industries agroalimentaires	7,65	Transports, postes et télécommunications	1,58
Textiles	9,82	Electricité, gaz et eau	0,62
Métallurgie fonderie (Or)	13,38	BTP	1,37
Autres industries	1,74	Commerce, restaurants et hôtels	2,60
Electricité et eau	0,70		
Construction	5,93		

Sources: Nos calculs, données INSAE-Bénin et INSTAT-Mali

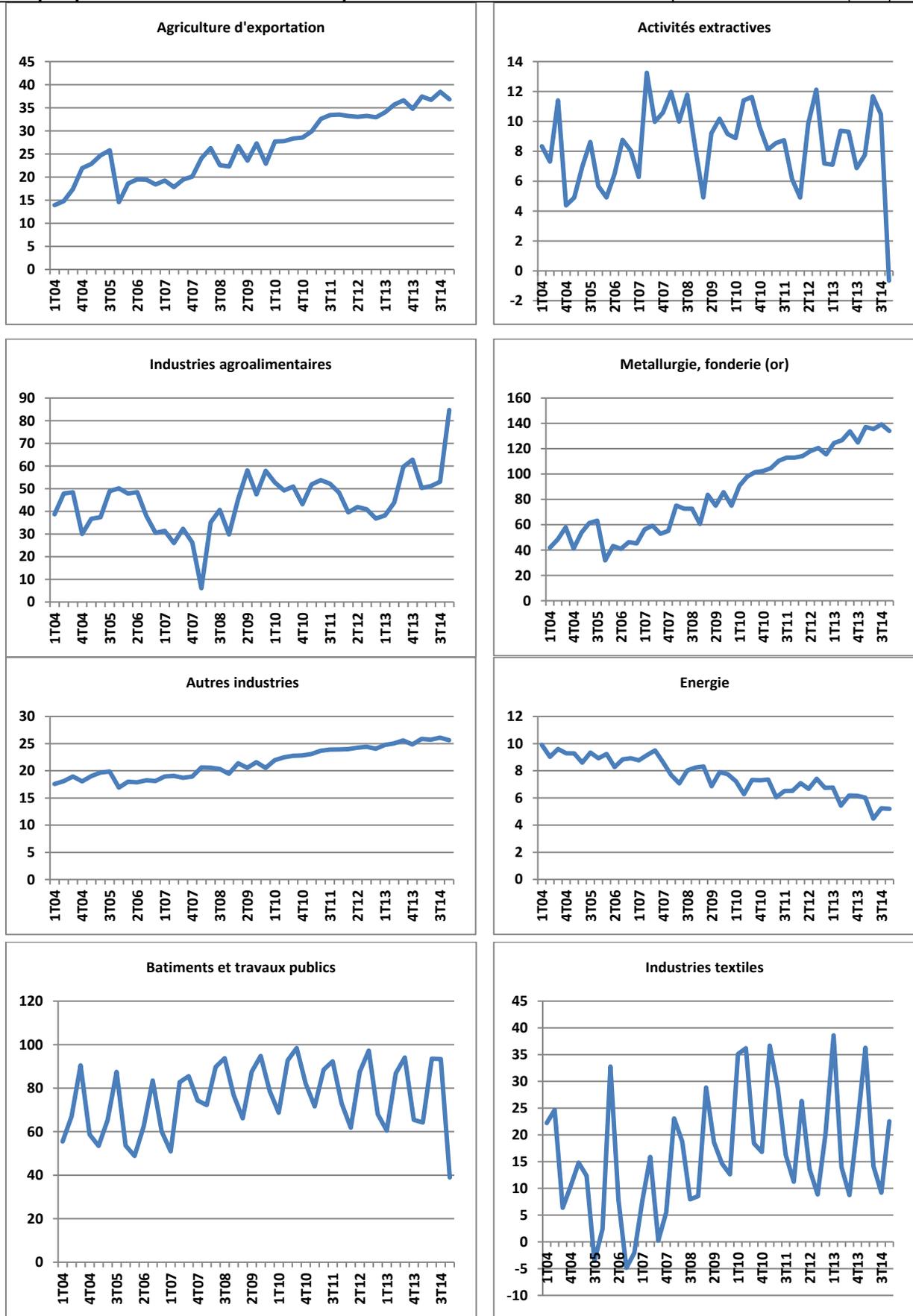
L'analyse des graphiques simultanées 1 et 2 couplés avec les erreurs quadratiques moyennes montre la bonne qualité des résultats obtenus sur les données du Bénin en comparaison à ceux du Mali, même si pour les deux branches « activités extractives » et « électricité et eau » les résultats sont plus satisfaisants.

**Graphique 3 : Evolution des valeurs ajoutées trimestrielles estimées sur la période 2001-2014 (Bénin)**



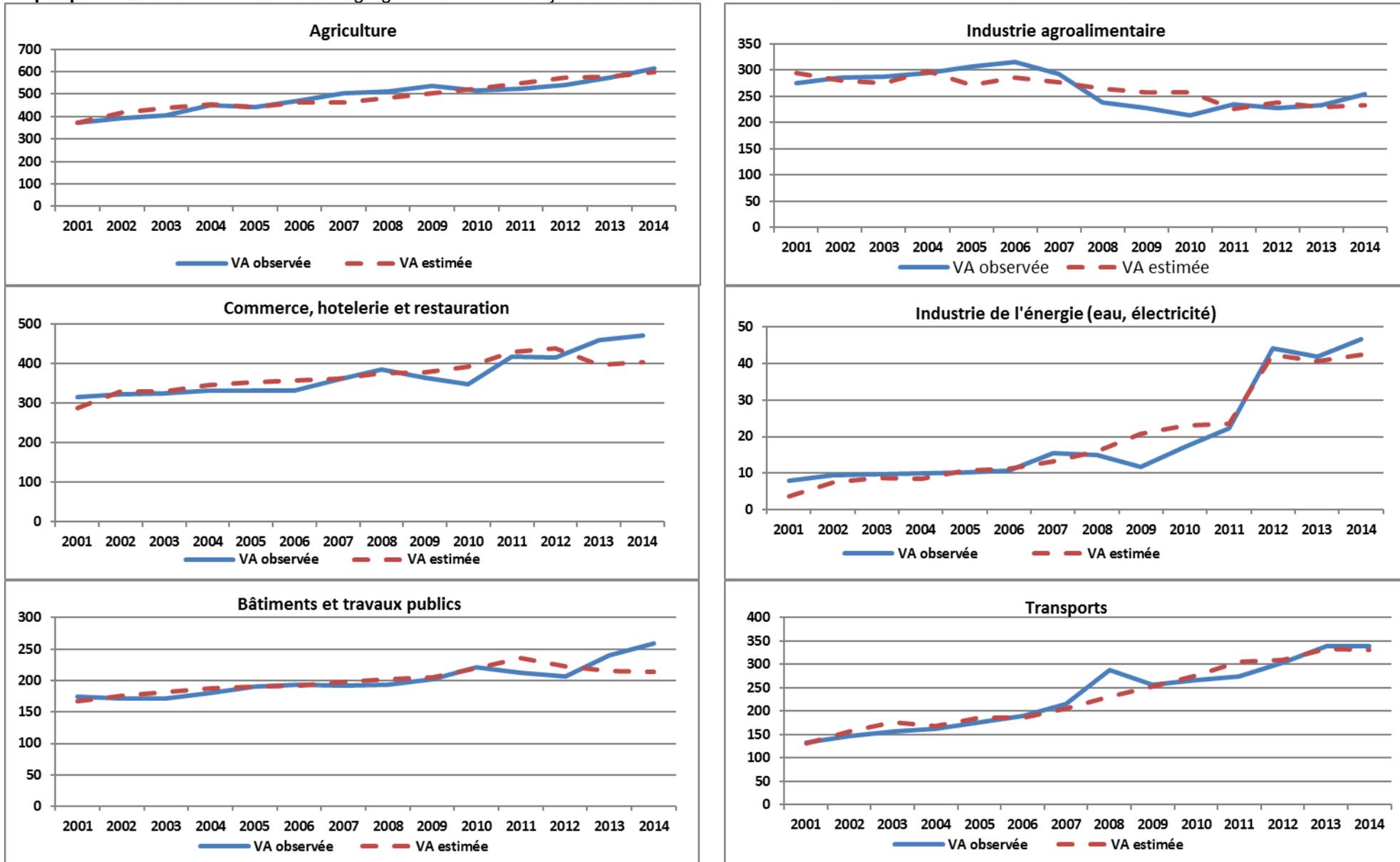
Sources: Les auteurs

**Graphique 4 : Evolution des valeurs ajoutées trimestrielles estimées sur la période 2004-2014 (Mali)**



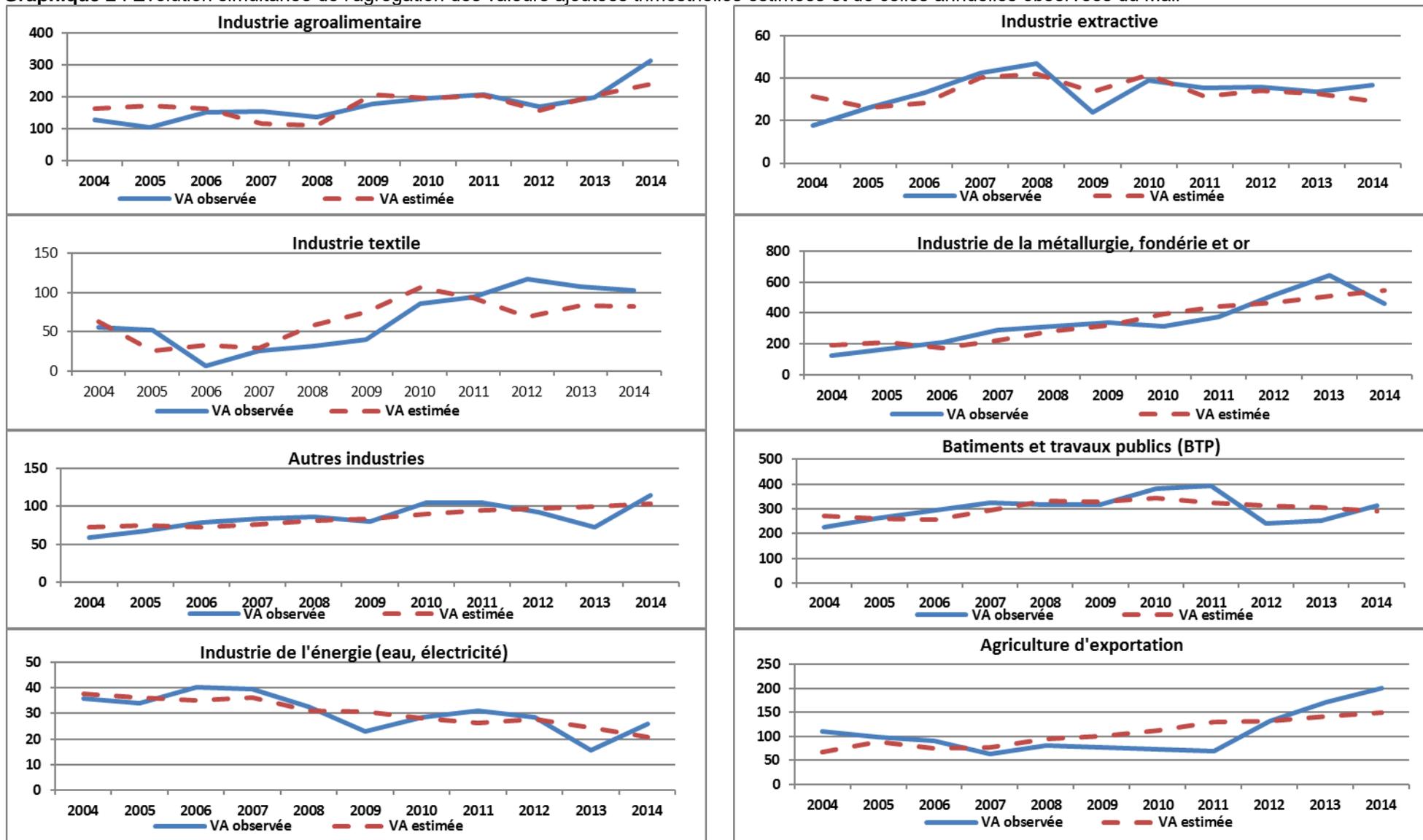
Sources: Les auteurs

**Graphique 1** : Evolution simultanée de l'agrégation des valeurs ajoutées trimestrielles estimées et de celles annuelles observées du Bénin



Sources: Les auteurs

**Graphique 2 :** Evolution simultanée de l'agrégation des valeurs ajoutées trimestrielles estimées et de celles annuelles observées du Mali



Sources: Les auteurs

### ***VI.3 Prédiction des valeurs ajoutées trimestrielles et annuelles de 2015***

On exploite l'information trimestrielle disponible sur les IHPI des années 2014 et 2015 pour prédire pour l'année 2015, les valeurs ajoutées trimestrielles inobservables et en déduire les valeurs ajoutées annuelles pour la même année. Les tableaux 6 et 7 (en annexes) présentent les prévisions ponctuelles et les intervalles de confiance.

Conformément à ce qui est attendu, les résultats mettent en relief le fait que les prévisions varient en fonction du trimestre avec une meilleure précision pour les valeurs ajoutées annuelles lorsque l'information trimestrielle disponible augmente. Il s'ensuit pour l'ensemble des branches que la longueur de l'intervalle de confiance se rétrécit lorsqu'on tend vers la fin de l'année de prévision.

Globalement, les résultats des prévisions sont satisfaisants si on se réfère au taux de recouvrement de la valeur ajoutée annuelle observée ; 83% des intervalles de confiance recouvrent la valeur ajoutée observée pour le Bénin en 2015 contre 75% pour le Mali.

**Tableau 6** : Prévisions trimestrielles des valeurs ajoutées sectorielles et annuelles pour le Bénin en 2015

Branche	Prévision ponctuelle et intervalle de confiance	Période relative à l'info disponible			
		T1	T2	T3	T4
<b>Agriculture</b>	<b>Valeur ajoutée annuelle</b>	<b>550,0</b>	<b>550,0</b>	<b>550,0</b>	<b>550,0</b>
	Prévision ponctuelle	593,52	593,74	593,47	593,44
	Intervalle de confiance à 95 %				
	borne inférieure	540,81	539,16	541,11	541,26
	borne supérieure	646,22	648,32	645,83	645,63
<b>Industries Agro-alimentaires</b>	<b>Valeur ajoutée observée</b>	<b>254,9</b>	<b>254,9</b>	<b>254,9</b>	<b>254,9</b>
	Prévision ponctuelle	226,77	227,39	234,64	234,03
	Intervalle de confiance à 95 %				
	borne inférieure	175,01	175,72	183,43	182,82
	borne supérieure	278,52	279,07	285,85	285,25
<b>Transports, postes et télécommunications</b>	<b>Valeur ajoutée observée</b>	<b>352,5</b>	<b>352,5</b>	<b>352,5</b>	<b>352,5</b>
	Prévision ponctuelle	353,05	359,24	351,64	350,89
	Intervalle de confiance à 95 %				
	borne inférieure	307,59	312,16	306,48	305,88
	borne supérieure	398,52	406,32	396,81	395,91
<b>Electricité, gaz et eau</b>	<b>Valeur ajoutée observée</b>	<b>40,2</b>	<b>40,2</b>	<b>40,2</b>	<b>40,2</b>
	Prévision ponctuelle	48,94	48,70	48,06	47,70
	Intervalle de confiance à 95 %				
	borne inférieure	40,46	40,22	39,60	39,25
	borne supérieure	57,42	57,17	56,52	56,16
<b>BTP**</b>	<b>Valeur ajoutée observée</b>	<b>280,8</b>	<b>280,8</b>	<b>280,8</b>	<b>280,8</b>
	Prévision ponctuelle	212,91	216,19	214,76	214,70
	Intervalle de confiance à 95 %				
	borne inférieure	177,52	180,08	179,03	178,99
	borne supérieure	248,30	252,29	250,49	250,42
<b>Commerce, restaurants et hôtels</b>	<b>Valeur ajoutée observée</b>	<b>473,0</b>	<b>473,0</b>	<b>473,0</b>	<b>473,0</b>
	Prévision ponctuelle	394,58	396,70	395,77	395,74
	Intervalle de confiance à 95 %				
	borne inférieure	328,14	328,92	328,71	328,69
	borne supérieure	461,01	464,48	462,84	462,78

Source: Nos calculs, données INSAE, Bénin; Période 2001-2014 ; (\*\*) valeur observée en dehors de l'intervalle de confiance

**Tableau 7 : Prévisions trimestrielles des valeurs ajoutées sectorielles et annuelles pour le Mali en 2015**

Branche	Période	Période relative à l'info disponible			
		T1	T2	T3	T4
<b>Agriculture d'exportation</b>					
	<b>Valeur ajoutée observée</b>	<b>188,72</b>	<b>188,72</b>	<b>188,72</b>	<b>188,72</b>
	Prévision ponctuelle	155,36	155,28	155,45	155,70
	Intervalle de confiance à 95 %				
	borne inférieure	76,29	76,08	76,54	77,04
	borne supérieure	234,42	234,49	234,36	234,37
<b>Activités extractives</b>					
	<b>Valeur ajoutée observée</b>	<b>38,6</b>	<b>38,6</b>	<b>38,6</b>	<b>38,6</b>
	Prévision ponctuelle	30,50	27,21	25,77	29,19
	Intervalle de confiance à 95 %				
	borne inférieure	16,68	12,94	11,47	15,74
	borne supérieure	44,33	41,48	40,08	42,65
<b>Industries agroalimentaires**</b>					
	<b>Valeur ajoutée observée</b>	<b>285,35</b>	<b>285,35</b>	<b>285,35</b>	<b>285,35</b>
	Prévision ponctuelle	145,80	149,37	152,25	156,64
	Intervalle de confiance à 95 %				
	borne inférieure	59,74	63,28	66,14	70,58
	borne supérieure	231,86	235,45	238,36	242,71
<b>Textiles *</b>					
	<b>Valeur ajoutée observée</b>	<b>93,65</b>	<b>93,65</b>	<b>93,65</b>	<b>93,65</b>
	Prévision ponctuelle	36,66	42,92	45,18	47,41
	Intervalle de confiance à 95 %				
	borne inférieure	-20,12	-12,20	-9,79	-7,57
	borne supérieure	93,43	98,03	100,14	102,40
<b>Métallurgie fonderie (Or)</b>					
	<b>Valeur ajoutée observée</b>	<b>471,79</b>	<b>471,79</b>	<b>471,79</b>	<b>471,79</b>
	Prévision ponctuelle	563,53	561,92	565,67	571,44
	Intervalle de confiance à 95 %				
	borne inférieure	398,93	397,02	401,39	407,66
	borne supérieure	728,14	726,82	729,95	735,21
<b>Autres industries</b>					
	<b>Valeur ajoutée observée</b>	<b>121,92</b>	<b>121,92</b>	<b>121,92</b>	<b>121,92</b>
	Prévision ponctuelle	104,97	104,85	105,13	105,56
	Intervalle de confiance à 95 %				
	borne inférieure	75,10	74,93	75,32	75,84
	borne supérieure	134,84	134,78	134,94	135,27
<b>Electricité et eau *</b>					
	<b>Valeur ajoutée observée</b>	<b>27,91</b>	<b>27,91</b>	<b>27,91</b>	<b>27,91</b>
	Prévision ponctuelle	22,88	19,96	18,04	16,37
	Intervalle de confiance à 95 %				
	borne inférieure	11,06	7,98	6,11	4,55
	borne supérieure	33,94	27,94	24,15	20,92
<b>Construction</b>					
	<b>Valeur ajoutée observée</b>	<b>338,74</b>	<b>338,74</b>	<b>338,74</b>	<b>338,74</b>
	Prévision ponctuelle	302,64	291,80	287,06	298,32
	Intervalle de confiance à 95 %				
	borne inférieure	208,67	194,81	189,81	206,85
	borne supérieure	511,32	486,61	476,87	505,17

Source: Source: Nos calculs, données INSTAT, Mali; Période 2004-2014 ; (\*) valeur observée en dehors de l'intervalle de confiance par moment ; (\*\*) valeur observée en dehors de l'intervalle de confiance

**Tableau 8 : Tests de Cointegration fondée sur l'approche GLS de Perron et Rodriguez (2016)**

Benin		MALI	
Branche	Statistique Z	Branche	Statistique Z
<b>Agriculture</b>	-11,6193***	<b>Agriculture d'exportation</b>	-8,72***
<b>Industries Agro-alimentaires</b>	-56,1189***	<b>Activités extractives Industries agroalimentaires</b>	-4,2139***
<b>Transports, postes et télécommunications</b>	-2,3798***	<b>Textiles</b>	-11,9323***
<b>Electricité, gaz et eau</b>	-2,9941***	<b>Métallurgie fonderie (Or)</b>	-3,085***
BTP	-3,5406***	<b>Autres industries</b>	-7,5213***
<b>Commerce, restaurants et hôtels</b>	-19,3959***	<b>Electricité et eau</b>	-4,5734***
		<b>Construction</b>	-3,4144***

(\*\*\*), (\*\*), (\*), significativité à 5%, à 7,5% et à 10%.

#### **VI.4 Conclusion tirée des cas pratiques et perspectives**

Cette étude a fait le choix d'un modèle à deux équations intégrant la dynamique saisonnière des indicateurs infra-annuels pour estimer et prédire les valeurs ajoutées trimestrielles inobservables et en déduire leur niveau annuel sur la base de l'information trimestrielle disponible.

Le modèle d'état caractérisant la dynamique de la variable infra-annuelle observable est un SAR(1), processus autorégressif saisonnier d'ordre 1. Il est couplé au modèle de mesure proposé pour les données annuelles observables, lequel est d'une portée essentiellement pratique puisque l'inférence sur les paramètres d'intérêt est fondée sur la loi normale standard pour deux d'entre eux et sur une distribution non standard dont les queues de distributions ont été tabulées par simulation de Monte Carlo pour le paramètre lié à la variation de l'indicateur infra-annuel.

Les paramètres d'intérêt sont en général très significatifs. L'agrégation annuelle des estimations trimestrielles des valeurs ajoutées comparée aux données observées conduit parfois à des erreurs quadratiques moyennes élevées, notamment lorsque le nombre d'observations est faible (autour de 10 pour le Mali). Cependant, les prévisions sont globalement satisfaisantes puisque plus de 80% des intervalles de confiance recouvrent la valeur ajoutée observée pour le Bénin en 2015 contre 75% pour le Mali.

Les résultats obtenus permettent de tirer comme enseignement majeur l'inadéquation pour certaines branches de l'indice harmonisé de production industrielle (IHPI) comme déterminant principal de la valeur ajoutée avec pour conséquence le choix de l'IHPI d'une autre branche ou de l'IPHI global. Il est donc important qu'une veille soit assurée pour l'efficacité du dispositif de production en mettant l'accent sur, notamment, la mise à jour de l'échantillonnage des entreprises de la branche concernée et la prise en compte du secteur informel.

En perspective, l'application de la méthodologie à l'ensemble des branches d'activité d'une économie permettra de prédire les PIB trimestriels pour en déduire les prévisions de taux de croissance.

Par ailleurs, l'extension de l'équation d'état à d'autres variables exogènes (prix des matières premières, prix du baril du pétrole, crise financière, etc.) pourrait être une voie d'approfondissement de la méthodologie au sens où elle permettra d'apprécier l'impact des chocs exogènes sur la croissance.

## Références bibliographiques

**AFRISTAT (2011)**, manuel méthodologique de production des comptes trimestriels dans les Etats d'Afrique au sud du Sahara, Séries Méthodes N°11.

**BCEAO (2014)**, Estimation du PIB trimestriel de l'UEMOA : approche méthodologique et résultats obtenus, Dakar (Sénégal).

**Dickey, D.A. and W.A. Fuller, (1979)**, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association* 74, 4277-431.

**Dickey, D.A., D.P. Hasza, and W.A. Fuller, (1984)**, Testing for unit roots in seasonal time series, *Journal of the American Statistical Association* 79, 3555-367.

**Elliott, G., T. Rothenberg, et J. H. Stock (1996)**, Efficient Tests for an Autoregressive Unit Root, *Econometrica*, 64,813-836.

**Fuller, W.A., (1976)**, Introduction of statistical time series (Wiley, New York, NY).

**Ghysels et al (1994)**, Testing for unit roots in seasonal time series: Some theoretical extensions and a Monte Carlo investigation\*, *Journal of Econometrics* 62 (1994) 415-442. North-Holland.

**Hjalmarsson, E. and P. Österholm (2007)**, Testing for Cointegration Using the Johansen Methodology when Variables are Near-Integrated, International Monetary Found WP/07/141.

**Hylleberg, S., R.F. Engle, C.W.J. Granger, and B.S. Yoo, (1990)**, Seasonal integration and cointegration, *Journal of Econometrics* 44, 215-238.

**INS-CAMEROUN (2012)**, Les comptes nationaux trimestriels : document de méthodologie, Yaoundé (Cameroun).

**Jan Kmenta (1991)** Latent variables in econometrics, Wiley Online Library Statistica Neerlandica Volume 45, Issue 2.

**Jansson, M. (2002)**, Consistent covariance matrix estimation for linear processes. *Econometric Theory*, 18,2002, 1449-1459.

**Lewis, J. B. and D. A. Linzer (2005)**, Estimating Regression Models in Which the Dependent Variable Is Based on Estimates, *Political Analysis* (2005), 13:345–364.

**Perron, P et G. Rodriguez (2016)**, Residuals-based Tests for Cointegration with GLS Detrended Data, *Econometrics Journal*, vol 19, pp. 84-111.

**Phillips, P.C.B. and B.E. Hansen (1990)**, Statistical inference in instrumental variables regression with I(1) variables. *Review of Economic Studies* 57, 99–125.

**Phillips, P.C.B. and P. Perron (1988)**, Testing for Unit Roots in Time Series Regression, *Biometrika*, 75, 335-346.

**R.F. Ngampana, I. Sory et C. Vodounou (2017)**, Estimation et prévision de valeurs ajoutées trimestrielles inobservables en présence de saisonnalité, STATECO, Numéro 111-2017.

## Annexes

### Annexes 1 : estimateurs des moindres carrés ordinaires de $\alpha$ , $\beta$ et $\gamma$ et leurs propriétés asymptotiques

**Proposition 1** : Soient les processus univariés  $\{y_n\}$  et  $\{X_t\}$  définis par les équations (1) et (2) et  $\{Q_n^S\}$  le processus annualisé de  $\{X_t\}$  tel que  $Q_n^S = \sum_{i=1}^S X_{S(n-1)+i}$ . Soient  $\bar{y} = N^{-1} \sum_{n=1}^N y_n$  et  $\bar{Q}^S = N^{-1} \sum_{n=1}^N Q_n^S$  pour  $n = 1, \dots, N$ . Sous l'hypothèse de faible stationnarité de  $\omega_n = (u_n, v_n)'$ , on pose  $\varpi_n = L(u_n, v_n)' = (\varepsilon_{u,v_n}, v_n)'$  où  $L = \begin{bmatrix} 1 & -\omega_{12}\omega_{22}^{-1} \\ 0 & 1 \end{bmatrix}$ . Si  $S_{[Nr]} = \sum_{n=1}^{[Nr]} \varpi_n$ , on a lorsque  $N \rightarrow \infty, N^{-1/2} S_{[Nr]} \Rightarrow B(r) = (B_1(r), B_2(r))' = (\varpi_{11}^{1/2} W_1(r), \omega_{22}^{1/2} W_2(r))$  où  $\varpi_{11} = \omega_{11} - \omega_{12}^2 \omega_{22}^{-1}$ ,  $\Omega = 2\pi f_{\omega_t}(0) = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{bmatrix}$  et  $W(r) = (W_1(r), W_2(r))'$  est un vecteur de dimension 2 de mouvements browniens standards indépendants. Sous ces hypothèses, on a:

$$i) Q_n^S = \exp(c/N) Q_{n-1}^S + v_n$$

$$ii) N^{-1/2} (Q_t^S - \bar{Q}^S) \Rightarrow \omega_{22}^{1/2} J_c^\mu(r) \quad \text{avec} \quad J_c^\mu(r) = J_c(r) - \int_0^1 J_c(s) ds \quad \text{et}$$

$$J_c(r) = \int_0^r \exp(c(r-s)) dW_2(s)$$

$$iii) N^{-2} \sum_{n=1}^N (Q_n^S - \bar{Q}^S)^2 \Rightarrow \omega_{22} \int_0^1 [J_c^\mu(r)]^2 dr$$

$$iv) N^{-1} \sum_{t=1}^T (Q_n^S - \bar{Q}^S) \varepsilon_{u,v_n} \Rightarrow \omega_{22}^{1/2} \varpi_{11}^{-1/2} \int_0^1 J_c^\mu(r) dW_1(r).$$

La preuve de cette proposition est établie sans peine. Elle est néanmoins présentée dans les travaux de Janssen(2002) et de façon particulière pour le cas où  $c = 0$  dans Phillips et Hansen (1990).

**Proposition 2** : L'estimateur  $\hat{\alpha}$  des moindres carrés ordinaires du paramètre  $\alpha$  est convergent et les estimateurs  $\hat{\beta}$  et  $\hat{\gamma}$  de  $\beta$  et  $\gamma$  sont hyperconvergents au sens où d'une part,  $\sqrt{N}(\hat{\alpha} - \alpha_0)$  et  $N(\hat{\beta} - \beta_0)$  sont asymptotiquement gaussiens et  $N(\hat{\gamma} - \gamma_0)$  converge vers une loi non standard et, d'autre part,  $(\hat{\beta} - \beta_0) = o_p(1)$ ,  $(\hat{\alpha} - \alpha_0) = o_p(1)$  et  $(\hat{\gamma} - \gamma_0) = o_p(1)$ . Les expressions des distributions asymptotiques sont données respectivement par:

$$N(\hat{\beta} - \beta_0) \Rightarrow \frac{\varpi_{11}^{1/2} \int_0^1 J_c^\mu(r) dW_1(r)}{\omega_{22}^{1/2} \int_0^1 [J_c^\mu(r)]^2 dr} \quad \sqrt{N}(\hat{\alpha} - \alpha_0) \Rightarrow -\varpi_{11}^{-1/2} \int_0^1 J_c(s) ds \frac{\int_0^1 J_c^\mu(r) dW_1(r)}{\int_0^1 [J_c^\mu(r)]^2 dr} \quad (14)$$

$$\text{et } N(\hat{\gamma} - \gamma_0) \Rightarrow -\frac{(\omega_{11}\omega_{22} - \omega_{12}^2\omega_{22}^{-1})^{1/2}}{\sigma_v^2} \frac{\int_0^1 J_c^\mu(r) dW_2(r)}{\int_0^1 [J_c^\mu(r)]^2 dr} \int_0^1 J_c^\mu(r) dW_1(r) \quad (15)$$

Les t-statistiques  $t_{\hat{\beta}}$  et  $t_{\hat{\alpha}}$  associés aux estimateurs  $\hat{\beta}$  et  $\hat{\alpha}$  sont asymptotiquement gaussiens. La loi asymptotique identique des t-statistiques est donnée par

$$\frac{\hat{\sigma}_{\varepsilon_{u,v_n}}}{\hat{\varpi}_{11}^{1/2}} t_{\hat{\beta}} \Rightarrow \frac{\int_0^1 J_c^\mu(r) dW_1(r)}{\left[ \int_0^1 [J_c^\mu(r)]^2 dr \right]^{1/2}} \equiv N(0,1) \quad (16)$$

Celle du t-statistique de  $\hat{\gamma}$  est donnée par

$$\frac{\hat{\sigma}_{\varepsilon u.v_n}}{\hat{\omega}_{11}^{1/2}} \frac{\hat{\sigma}_v}{\omega_{22}^{1/2}} N^{1/2} t_{\hat{\gamma}} \Rightarrow - \frac{\int_0^1 J_c^\mu(r) dW_2(r)}{\left[ \int_0^1 [J_c^\mu(r)]^2 dr \right]^{1/2}} \frac{\int_0^1 J_c^\mu(r) dW_1(r)}{\left[ \int_0^1 [J_c^\mu(r)]^2 dr \right]^{1/2}} \quad (17)$$

où  $\hat{\sigma}_{\varepsilon u.v_n}^2 = N^{-1} \sum_{n=1}^N \hat{\varepsilon}_{u.v_n}^2$

et  $\hat{\omega}_{11} = 2\pi \hat{f}_{\varepsilon u.v_n}(0) = N^{-1} \sum_{n=1}^N \hat{\varepsilon}_{u.v_n}^2 + 2N^{-1} \sum_{\tau=1}^l \sum_{n=1+\tau}^N \left(1 - \frac{\tau}{1+l}\right) \hat{\varepsilon}_{u.v_\tau} \hat{\varepsilon}_{u.v_{n-\tau}}$

**Remarque :** i) Si  $Q_n^S$  était exogène, alors le paramètre  $\beta$  pourrait être estimé par la régression (5)

sans l'extension à la variable  $\Delta_c Q_n^S$  et  $\frac{\hat{\sigma}_{\varepsilon u.v_n}}{\hat{\omega}_{11}^{1/2}} t_{\hat{\beta}} \Rightarrow N(0,1)$ . Si de plus  $u_n$  était i.i.d $(0, \sigma_u^2)$  alors

$t_{\hat{\beta}} \Rightarrow N(0,1)$ .

ii) Si  $\omega_n = (u_n, v_n)'$  est i.i.d $(0, \Sigma)$  alors  $\Omega = 2\pi f_{\omega_t}(0) = \Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix}$ .

iii) La distribution asymptotique de la statistique pi (votale  $\frac{\hat{\sigma}_{\varepsilon u.v_n}}{\hat{\omega}_{11}^{1/2}} \frac{\hat{\sigma}_v}{\omega_{22}^{1/2}} N^{1/2} t_{\hat{\gamma}}$ ) donnée par

(10) est non standard. Elle est le produit de la normale centrée réduite  $N(0,1)$

$\equiv \frac{\int_0^1 J_c^\mu(r) dW_1(r)}{\left[ \int_0^1 [J_c^\mu(r)]^2 dr \right]^{1/2}}$  et celle de Dickey-Fuller (1979), tabulée déjà par Fuller (1976) dans

le cadre des tests de racine unitaire lorsque  $c = 0$ . L'extension de cette dernière à des valeurs non nulles de  $c$  a été tabulée par Phillips et Perron (1988). Il est important de noter que l'utilisation du  $t_{\hat{\gamma}}$  sans le facteur multiplicateur  $N^{1/2}$  conduirait au non rejet le

plus souvent de l'hypothèse nulle  $H_0: \gamma = \gamma_0$ .

iv) Le coefficient de détermination  $R^2$  converge vers 1 lorsque  $N$  tend vers l'infini.

### Preuve de la proposition 2 :

Soit la régression augmentée définie par (5)

$$y_n = \alpha + \beta Q_n^S + \gamma \Delta_c Q_n^S + \eta_n \quad (18)$$

Les paramètres  $\beta$  et  $\gamma$  sont estimés par OLS en considérant l'équation

$$y_n - \bar{y} = \beta (Q_n^S - \bar{Q}^S) + \gamma (\Delta_c Q_n^S - \overline{\Delta_c Q}^S) + \eta_n \quad (19)$$

et le paramètre  $\alpha$  est estimé par  $\hat{\alpha} = \bar{y} - \hat{\beta} \bar{Q}^S - \hat{\gamma} \overline{\Delta_c Q}^S$

On pose :  $\Delta = \Sigma (Q_n^S - \bar{Q}^S)^2 \Sigma (\Delta_c Q_n^S - \overline{\Delta_c Q}^S)^2 - \left[ \Sigma (Q_n^S - \bar{Q}^S) (\Delta_c Q_n^S - \overline{\Delta_c Q}^S) \right]^2$  et on note que  $v_n = \Delta_c Q_n^S$ .

Les estimateurs des moindres carrés ordinaires de  $\beta$  et  $\gamma$  sont donnés sous l'hypothèse nulle par :

$$\hat{\beta} - \beta_0 = \frac{\Sigma (Q_n^S - \bar{Q}^S) \varepsilon_{u.v_n} + \frac{\Sigma (Q_n^S - \bar{Q}^S) v_n}{\Sigma (v_n - \bar{v})^2} \Sigma v_n \varepsilon_{u.v_n}}{\Sigma \{ (Q_n^S - \bar{Q}^S) - \frac{\Sigma (Q_n^S - \bar{Q}^S) v_n}{\Sigma (v_n - \bar{v})^2} (v_n - \bar{v}) \}^2}$$

$$\text{et } \hat{\gamma} - \gamma_0 = \frac{-[\sum(Q_n^S - \bar{Q}^S)v_n][\sum(Q_n^S - \bar{Q}^S)\varepsilon_{u.v_n}] + \sum(Q_n^S - \bar{Q}^S)^2 \sum v_n \varepsilon_{u.v_n}}{\sum(Q_n^S - \bar{Q}^S)^2 \sum (v_n - \bar{v})^2 - [\sum(Q_n^S - \bar{Q}^S)v_n]^2}$$

Les processus  $v_n$  et  $\varepsilon_{u.v_n}$  ne sont pas corrélés. On en déduit lorsque  $N$  tend vers l'infini que

$$N(\hat{\beta} - \beta_0) \Rightarrow \frac{\omega_{11}^{1/2} \int_0^1 J_c^\mu(r) dW_1(r)}{\omega_{22}^{1/2} \int_0^1 [J_c^\mu(r)]^2 dr} \quad (20)$$

$$\text{et } N(\hat{\gamma} - \gamma_0) \Rightarrow \frac{\omega_{22}^{1/2} \omega_{11}^{1/2} \int_0^1 J_c^\mu(r) dW_1(r)}{\hat{\sigma}_v^2 \int_0^1 [J_c^\mu(r)]^2 dr} \int_0^1 J_c^\mu(r) dW_2(r) \quad (21)$$

d'où l'hyperconvergence de  $\hat{\beta}$  et de  $\hat{\gamma}$ ; c'est à dire :  $\hat{\beta} - \beta_0 \Rightarrow 0$  et  $\hat{\gamma} - \gamma_0 \Rightarrow 0$ .

La variance de  $\hat{\beta}$  est estimée par:  $\text{var} \hat{\beta} = \hat{\sigma}_{\varepsilon_{u.v_n}}^2 \frac{\sum (v_n - \bar{v})^2}{\sum(Q_n^S - \bar{Q}^S)^2 \sum (v_n - \bar{v})^2 - [\sum(Q_n^S - \bar{Q}^S)v_n]^2}$

On en déduit, sans peine que :  $N^2 \text{var} \hat{\beta} \Rightarrow \frac{\hat{\sigma}_{\varepsilon_{u.v_n}}^2}{\omega_{22} \int_0^1 [J_c^\mu(r)]^2 dr}$

Avec la même logique, on a:  $\text{var} \hat{\gamma} = \hat{\sigma}_{\varepsilon_{u.v_n}}^2 \frac{\sum(Q_n^S - \bar{Q}^S)^2}{\sum(Q_n^S - \bar{Q}^S)^2 \sum (v_n - \bar{v})^2 - [\sum(Q_n^S - \bar{Q}^S)v_n]^2}$  et

$$N \text{ var} \hat{\gamma} \Rightarrow \frac{\hat{\sigma}_{\varepsilon_{u.v_n}}^2}{\hat{\sigma}_v^2}$$

Par conséquent, on a par passage à la limite :  $t_{\hat{\beta}} \Rightarrow \frac{\omega_{11}^{1/2} \int_0^1 J_c^\mu(r) dW_1(r)}{\hat{\sigma}_{\varepsilon_{u.v_n}} \left[ \int_0^1 [J_c^\mu(r)]^2 dr \right]^{1/2}}$  (22)

et  $N^{1/2} t_{\hat{\gamma}} \Rightarrow -\frac{\omega_{11}^{1/2} \omega_{22}^{1/2} \int_0^1 J_c^\mu(r) dW_1(r)}{\hat{\sigma}_{\varepsilon_{u.v_n}} \hat{\sigma}_v} \frac{\int_0^1 J_c^\mu(r) dW_2(r)}{\left[ \int_0^1 [J_c^\mu(r)]^2 dr \right]^{1/2}}$  (23)

S'agissant de  $t_{\hat{\alpha}}$ , on a d'abord:  $\sqrt{N}(\hat{\alpha} - \alpha_0) = -N(\hat{\beta} - \beta_0)N^{-1/2}\bar{Q}^S + o_P(1)$  (24)

Comme  $N^{-1/2}\bar{Q}^S = N^{-3/2}\sum Q_n^S \Rightarrow \omega_{22}^{1/2} \int_0^1 J_c(s) ds$ , il s'ensuit que:

$$\sqrt{N}(\hat{\alpha} - \alpha_0) \Rightarrow -\omega_{11}^{1/2} \int_0^1 J_c(s) ds \frac{\int_0^1 J_c^\mu(r) dW_1(r)}{\int_0^1 [J_c^\mu(r)]^2 dr} \quad (25)$$

D'où  $\sqrt{N}(\hat{\alpha} - \alpha_0)$  suit asymptotiquement une loi normale conditionnellement à  $J_c^\mu(r)$ . L'estimateur des moindres carrés ordinaires de  $\alpha$  est convergent et on a :  $\hat{\alpha} - \alpha_0 \Rightarrow 0$

Pour la variance de  $\hat{\alpha}$ , on montre sans peine que  $N \text{ var} \hat{\alpha} \Rightarrow \hat{\sigma}_{\varepsilon_{u.v_n}}^2 \frac{(\int_0^1 J_c(s) ds)^2}{\int_0^1 [J_c^\mu(r)]^2 dr}$

$$t_{\hat{\alpha}} = \frac{\hat{\alpha} - \alpha_0}{[\text{var} \hat{\alpha}]^{1/2}} \Rightarrow -\frac{\omega_{11}^{1/2} \int_0^1 J_c^\mu(r) dW_1(r)}{\hat{\sigma}_{\varepsilon_{u.v_n}} \left[ \int_0^1 [J_c^\mu(r)]^2 dr \right]^{1/2}} \quad (26)$$

Il s'ensuit, pour raison de symétrie que  $\frac{\hat{\sigma}_{\varepsilon_{u.v_n}}}{\omega_{11}^{1/2}} t_{\hat{\beta}}$  et  $\frac{\hat{\sigma}_{\varepsilon_{u.v_n}}}{\omega_{11}^{1/2}} t_{\hat{\alpha}}$  ont la même loi asymptotique  $N(0,1)$ .

**Annexes 2 : Queues de distribution de la loi asymptotique de la statistique de Student**

$$\hat{\gamma} : - \frac{\int_0^1 J_c^H(r) dW_1(r)}{\left[ \int_0^1 [J_c^H(r)]^2 dr \right]^{1/2}} \frac{\int_0^1 J_c^H(r) dW_2(r)}{\left[ \int_0^1 [J_c^H(r)]^2 dr \right]^{1/2}}$$

**normalisée de**

	0,010	0,025	0,050	0,100	0,500	0,900	0,950	0,975	0,990
c=-15	-3,191	-2,261	-1,618	-1,050	-0,003	1,082	1,623	2,213	3,027
c=-10	-2,939	-2,153	-1,605	-1,025	-0,001	1,026	1,592	2,120	2,931
c=-5	-3,038	-2,219	-1,567	-1,022	0,001	0,994	1,502	2,125	2,873
c=-2	-3,029	-2,161	-1,593	-1,035	-0,004	0,999	1,529	2,167	2,946
c=-1	-3,062	-2,254	-1,630	-1,023	0,001	1,059	1,634	2,213	3,046
c=0	-2,939	-2,177	-1,623	-1,048	0,001	1,002	1,581	2,170	3,030
c=1	-3,044	-2,191	-1,600	-1,052	-0,001	1,030	1,615	2,183	2,959
c=2	-2,929	-2,145	-1,574	-1,000	0,003	1,039	1,596	2,116	2,909
c=5	-2,968	-2,116	-1,544	-1,005	-0,001	1,020	1,518	2,086	2,840
c=10	-3,017	-2,267	-1,622	-1,070	-0,005	1,007	1,553	2,169	3,072
c=15	-2,906	-2,086	-1,550	-0,992	0,003	1,035	1,577	2,170	3,002
N(0,1)*	-2,326	-1,960	-1,645	-1,282	0,000	1,282	1,645	1,960	2,326

Source : Nos simulations sur la base de 10000 répliques de Monte Carlo ; (\*) tirée des tables existantes pour comparaison

**Annexes 3 : Coefficient de détermination  $R^2$**

Le coefficient de détermination s'écrit

$$R^2 = \frac{\sum_{i=1}^N (\hat{y}_n - \bar{y})^2}{\sum_{i=1}^N (y_n - \bar{y})^2} = \frac{\sum_{i=1}^N \{ \hat{\beta}(Q_n^S - \bar{Q}^S) + \hat{\gamma}(\Delta_c Q_n^S - \Delta_c \bar{Q}^S) \}^2}{\sum_{i=1}^N \{ \beta_0(Q_n^S - \bar{Q}^S) + \gamma_0(\Delta_c Q_n^S - \Delta_c \bar{Q}^S) + \varepsilon_{u,v_n} \}^2}$$

On a :  $N^{-2} \sum_{i=1}^N (\hat{y}_n - \bar{y})^2 = (\hat{\beta})^2 N^{-2} \sum_{i=1}^N (Q_n^S - \bar{Q}^S)^2 + o(1)_P$

et  $N^{-2} \sum_{i=1}^N (y_n - \bar{y})^2 = (\beta_0)^2 N^{-2} \sum_{i=1}^N (Q_n^S - \bar{Q}^S)^2 + o(1)_P$

Lorsque  $N$  tend vers l'infini,  $\hat{\beta} \Rightarrow \beta_0$  sous  $H_0$  On en déduit que  $R^2 \Rightarrow 1$ .

D'un point de vue pratique  $R^2$  a peu d'intérêt sous l'hypothèse nulle. Hjalmarsson (2008) suggère d'étudier dans ce cas le comportement de  $R^2$  sous des alternatives locales.

## Annexes 4 : Tests de racine unitaire et de cointégration

### 1. Rappel des tests de racine unitaire classique (ADF, PP, KPSS)

Les tests usuels s'hypothèse nulle de non-stationnarité contre l'hypothèse alternative de stationnarité sont ceux de Dickey-Fuller (1981), Said-Dickey (1988), Phillips-Perron (1988) et leurs versions modifiés auxquels on rajoute les tests efficaces de Elliot et al (1997).

#### 1.1. Test de Dickey-Fuller (DF)

Soit  $y_t$  le processus généré par l'un des modèle suivants:

$$y_t = y_{t-1} + \varepsilon_t \quad (27)$$

où  $\varepsilon_t : iidN(0, \sigma^2)$  et  $t = 1, 2, 3, \dots, T$ .

Et les régressions

$$y_t = TD_t + \alpha y_{t-1} + u_t \quad (28)$$

où  $TD_t = 0, \mu$  ou  $\mu + \beta t$ .

On s'intéresse au test de l'hypothèse  $H_0 : \alpha = 1$  contre  $H_1 : \alpha < 1$ . L'estimateur des moindres carrés ordinaires de  $\alpha$  fondé sur (28) avec  $TD_t = 0$  est donné par

$$\hat{\alpha} = \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_{t-1}^2} \quad (29)$$

Dickey(1976) et Fuller(1976) considèrent les statistiques de test  $T \left( \hat{\alpha} - 1 \right)$  et

$t_{\hat{\alpha}} = \left( \hat{\alpha} - 1 \right) \left( \sum_{t=1}^T y_{t-1}^2 \right)^{1/2} / s$  avec  $s^2 = T^{-1} \sum_{t=1}^T \left( y_t - \hat{\alpha} y_{t-1} \right)^2$ . Sous  $H_0 : \alpha = 1$ ,

les distributions asymptotiques de ces statistiques sont données par:

$$T \left( \hat{\alpha} - 1 \right) \rightarrow \frac{\frac{1}{2} \left( [W(1)]^2 - 1 \right)}{\int_0^1 W^2(r) dr} = \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W^2(r) dr} \quad (30)$$

$$t_{\hat{\alpha}} \rightarrow \frac{\frac{1}{2} \left( [W(1)]^2 - 1 \right)}{\left[ \int_0^1 W^2(r) dr \right]^{1/2}} = \frac{\int_0^1 W(r) dW(r)}{\left[ \int_0^1 W^2(r) dr \right]^{1/2}} \quad (31)$$

où  $W(r)$  est le processus de Wiener standard.

Les distributions (30) et (31) obtenus par simulation de Monte Carlo sont non standards et tabulées dans Fuller(1976). Dans le cas où  $TD_t = \mu$  on remplace  $W(r)$  par  $W_*(r) = W(r) - \int_0^1 W(s) d(s)$ .

Lorsque  $TD_t = \mu + \beta t$ , on remplace  $W(r)$  par le résidu de sa projection sur  $(1, r)$  noté

$$W_{**}(r) = W(r) + (6r - 4) \int_0^1 W(s) d(s) - (12r - 6) \int_0^1 s W(s) d(s)$$

Les distributions asymptotiques des estimateurs des moindres carrés ordinaires de  $\mu$  et  $\beta$  sont données dans Philips et Perron (1987) et tabulées dans Dickey-Fuller (1981).

Dans la pratique les résidus obtenus dans (28), ou dans les régressions avec constante ou tendance sont en général corrélés et l'hypothèse de normalité n'est pas toujours satisfaite. Comme remède au problème de corrélation, Said et Dickey (1979) ont proposé d'introduire dans les régressions (28), un nombre suffisant de retards. Les tests ADF ci-après sont développés dans cette perspective.

### 1.2. Test de Dickey - Fuller augmenté (ADF)

On considère la régression

$$\Delta y_t = \eta + \gamma t + \rho y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-j} + u_t \quad (32)$$

L'hypothèse nulle s'écrit  $H_0 : \rho = (\alpha - 1) = 0$  contre l'hypothèse alternative de stationnarité  $H_1 : \rho < 0$ .

On suppose que dans (32),  $\eta = 0$  et  $\gamma = 0$ . On pose

$$x_t = (\Delta y_{t-1}, \Delta y_{t-2}, \dots, \Delta y_{t-k})'$$

Soit  $R_{ot}$  le résidu de la régression de  $\Delta y_t$  sur  $x_t$ ; puis par  $R_{1t}$  le résidu de la régression de  $y_{t-1}$  sur  $x_t$ . On a:

Les statistiques de test sont

$$t(ADF) = \frac{\rho}{\sigma_u \left( \sum_{t=1}^T R_{1t}^2 \right)^{-1/2}}$$

et

$$\frac{T \left( \alpha - 1 \right)}{1 - \sum_{i=1}^k c_i} = \frac{T \rho}{1 - \sum_{i=1}^k c_i}$$

Les distributions asymptotiques sont données dans Dickey et Fuller (1981) et Chang et Park (2001) et sont identiques à celles étudiées pour les tests DF et tabulées dans Fuller (1976).

Le problème pratique important rencontré dans la mise en oeuvre du test ADF est le choix du nombre de retard  $k$ . Ce choix est fondé sur les critères d'information tels que Akaike (AIC): 1973, Schwartz (BIC): 1978, et Hannan et Quinn (HQ): 1979 qui conduisent à un choix de  $k$  petit correspondant à l'argument de minimisation de tels critères.

$$IC = \ln \left( \sigma_k^2 \right) + (k + p + 1) C_T / T$$

où  $\sigma_k^2 = T^{-1} \sum_{t=1}^T e_{tk}^2$ ,  $k$  le nombre de retards et  $(p+1)$  le nombre de termes déterministes. AIC correspond à  $C_T = 2$ ; BIC correspond à  $C_T = \ln(T)$ ; et HQ à  $C_T = \ln(\ln(T))$ .  $C_T$  est le facteur de pénalité. Il vérifie  $C_T/T \rightarrow 0$  lorsque  $T \rightarrow \infty$ . Soit  $k_{IC}$  le choix de  $k$  résultant de la minimisation de  $IC$ . on a:

$$k_{IC} = \arg \min_k \left[ \ln \left( \sigma_k^2 \right) + (k + p + 1) C_T / T \right]$$

L'utilisation des critères d'information conduit généralement au choix de modèle parcimonieux ( $k_{IC}$  faible).

L'approche alternative de Ng et Perron (1995) fondée sur un choix de  $k_{\max}$  choisi a priori mais dépendant du nombre d'observations. Le critère asymptotique du choix de  $k_{\max}$  est  $k_{\max} = o(T^{1/3})$ ; ce qui signifie que  $k^3/T \rightarrow 0$  lorsque  $k, T \rightarrow \infty$  [Berk(1974) et Said et Dickey(1984)]. On peut aussi utiliser le critère suggéré par Schwert(1989) à savoir  $k_{\max} = \left[ 12 \left( \frac{T}{100} \right)^{1/4} \right]$ .

Une fois choisi  $k_{\max}$ , la sélection du  $k$  optimal dans la régression (32) est fondée sur une procédure descendante qui consiste à estimer par moindres carrés ordinaires les coefficients  $c_j (j=1, 2, \dots, k_{\max})$ ,  $\rho, \eta, \gamma$  puis à tester la significativité du dernier coefficient  $c_{k_{\max}}$ . Si  $c_{k_{\max}} \neq 0$  alors  $k = k_{\max}$ . Sinon, on diminue  $k_{\max}$  d'une unité. On estime de nouveau les coefficients et on teste la significativité de  $c_{k_{\max}-1}$ . Le  $k$  optimal correspond à la valeur de  $k$  pour laquelle le dernier coefficient  $c_k$  est significatif. Les simulations menées par Ng et Perron (1995) ont montré qu'avec cette procédure, les statistiques de test ont de bonnes propriétés: niveau de test stable et perte de puissance minimisée.

## 1.2. Tests de Phillips et Perron (PP)

Les tests de Phillips et Perron(1988) diffèrent de ceux de Dickey-Fuller augmentés. Ils corrigent directement les statistiques de tests de racine unitaire en prenant en compte le problème de corrélation et d'hétéroscédasticité qui pourrait subsister dans le terme aléatoire. Ils considèrent à cet effet, le processus générateur des données:

$$\begin{aligned} y_t &= \alpha y_{t-1} + u_t \\ \alpha &= 1 \end{aligned} \quad (33)$$

où  $u_t$  vérifie les conditions générales suivantes:

- i)  $E(u_t) = 0$  pour tout  $t$ ;
- ii)  $\sup_t E|u_t|^{\beta+\varepsilon} < \infty$  pour  $\beta > 2$  et  $\varepsilon > 0$ ;
- iii)  $\sigma^2 = \lim_{T \rightarrow \infty} E(T^{-1} S_T^2)$  existe et  $\sigma^2 > 0$  avec  $S_T = \sum_{t=1}^T u_t$
- iv)  $\{u_t\}$  est fortement mélangeant avec pour coefficients  $\alpha_m$  vérifiant  $\sum_{m=1}^{\infty} \alpha_m^{1-2/\beta} < \infty$ .

Ces conditions incluent des mécanismes de génération des données plus généraux comme les processus ARMA et autorisent l'hétéroscédasticité dans  $u_t$ . Si  $u_t$  est faiblement stationnaire de fonction de densité spectrale  $f_u(\omega)$  alors iii) découle de ii) et de iv).

On a dans ce cas  $\sigma_u^2 = \lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T E[u_t^2]$  et

$$\sigma^2 = 2\pi f_u(0) = E(u_1^2) + 2 \sum_{k=2}^{\infty} E(u_1 u_k) = \lim_{T \rightarrow \infty} \sum_{t=1}^T [T^{-1} S_T^2]$$

On considère les régressions

$$\Delta y_t = \beta TD_t + \rho y_{t-1} + u_t \quad (34)$$

où  $TD_t = 0$  (absence de constante et de tendance),  $TD_t = 1$  (modèle avec constante) et  $TD_t = (1, t)$  (modèle avec constante et tendance).

Sous ces hypothèses sur  $u_t$ , les statistiques  $T \begin{pmatrix} \rho \\ \alpha - 1 \end{pmatrix}$  et  $t_{\alpha}$  ont des comportements asymptotiques différents de ceux tabulés dans Fuller (1976). En effet, on a pour  $TD_t = 0$  :

$$T \begin{pmatrix} \rho \\ \alpha - 1 \end{pmatrix} \rightarrow \frac{\int_0^1 W(r) dW(r) + \lambda}{\int_0^1 W^2(r) dr} \quad (35)$$

$$t_{\alpha} \rightarrow \left( \frac{\sigma}{\sigma_u} \right) \frac{\int_0^1 W(r) dW(r) + \lambda}{\left[ \int_0^1 W^2(r) dr \right]^{1/2}} \quad (36)$$

où  $\lambda = \left( 1 - \frac{\sigma_u^2}{\sigma^2} \right) / 2$ . Pour  $TD_t = 1$  ou  $TD_t = (1, t)$ ,

les distributions  $\left( \int_0^1 W(r) dW(r), \int_0^1 W^2(r) dr \right)$  sont remplacées respectivement

par  $\left( \int_0^1 W_*(r) dW(r), \int_0^1 W_*^2(r) dr \right)$  et  $\left( \int_0^1 W_{**}(r) dW(r), \int_0^1 W_{**}^2(r) dr \right)$ . Comme on le constate,

les distributions ci-dessus dépendent des paramètres de nuisance  $\sigma_u^2$ ,  $\sigma^2$  et  $\lambda$ .

**Phillips et Perron (1988) proposent une modification des statistiques usuelles de test afin d'obtenir les mêmes distributions asymptotiques que ceux des tests DF.** Ainsi  $Z_{\alpha}$  et  $Z(t_{\alpha})$

modifient respectivement  $T \begin{pmatrix} \rho \\ \alpha - 1 \end{pmatrix}$  et  $t_{\alpha}$ . Elles sont définies par

$$Z_{\alpha} = T \begin{pmatrix} \rho \\ \alpha - 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} s^2 & -s^2 \\ s & s \end{pmatrix} \left( T^{-2} \sum_{t=1}^T y_{t-1}^2 \right)^{-1} \quad (37)$$

$$Z(t_\alpha) = \frac{s_u}{s} t_\alpha - \left( \frac{1}{2} \right) \left( \frac{s^2}{s} - \frac{s^2}{s} \right) \left( \frac{s^2}{s} T^{-2} \sum_{t=1}^T y_{t-1}^2 \right)^{-1/2} \quad (38)$$

où  $y_{t-1}$  est le résidu de la régression de  $y_{t-1}$  sur  $TD_t = 0, 1, (1, t)$ .  $s^2$  et  $s_u^2$  sont des estimateurs convergents de  $\sigma^2$  et  $\sigma_u^2$ . Ils sont estimés par

$$s_u^2 = T^{-1} \sum_{t=1}^T u_t^2$$

avec  $u_t = y_t - \beta TD_t - \alpha y_{t-1}$ ,  $\alpha = 1 + \rho$  lorsqu'on utilise la régression (34) et

$$s^2 = T^{-1} \sum_{t=1}^T u_t^2 + 2 \sum_{s=1}^l w_{sl} \left( T^{-1} \sum_{t=s+1}^T u_t u_{t-s} \right)$$

avec  $w_{sl} = 1 - s/(1+l)$ . Le paramètre de troncature  $l$  est choisi tel que  $l \rightarrow \infty$  et  $l^4/T \rightarrow 0$ .

**Remarque 4** Perron et Ng (1996, 1998) puis Stock (1999) proposent en suivant les travaux de Berk (1974) d'estimer  $s^2$  par une approche autorégressive fondée sur la régression (32) sous l'hypothèse nulle. On a ainsi

$$s_{AR}^2 = s_u^2 / \left( 1 - b(1) \right)^2 \quad (39)$$

avec  $b(1) = \sum_{j=1}^k b_j$ , les paramètres  $b_j$  étant estimés par la régression suivante

$$\Delta y_t = c + b_0 y_{t-1} + \sum_{j=1}^k b_j \Delta y_{t-j} + u_t \quad (40)$$

Les statistiques  $Z_\alpha$  et  $Z(t_\alpha)$  ont sous l'hypothèse nulle, respectivement des distributions asymptotiques identiques à celles des tests DF.

On en déduit que les statistiques  $Z_\alpha$ ,  $Z(t_\alpha)$ ,  $T \left( \alpha - 1 \right)$  et  $t_\alpha$  **ont les mêmes propriétés pour**

**des puissances asymptotiques locales.**

Au total, les tests ADF et PP présentent à distance finie des distorsions de niveau quand les erreurs aléatoires ont des corrélations sérielles. Ces distorsions sont d'autant plus importantes que les erreurs aléatoires suivent un MA (1) dont la racine est proche de -1.

#### Tests de racine unitaire modifiés (M tests)

Une réponse au problème de distorsion de niveau observé avec les tests PP est apportée par Perron et Ng (1996) qui proposent les modifications suivantes.

i) **La statistique  $MZ_\alpha$**  est une modification de la statistique  $Z_\alpha$ . Elle est définie par

$$MZ_\alpha = Z_\alpha + \frac{T}{2} \left( \alpha - 1 \right)^2 \quad (41)$$

Les statistiques  $MZ_\alpha$  et  $Z_\alpha$  sont asymptotiquement équivalentes: car  $\alpha - 1 = O_p(T^{-1})$  sous l'hypothèse nulle. Elles ont par conséquent la même distribution asymptotique.

ii) La statistique  $MZ(t_\alpha)$  est une modification de  $Z(t_\alpha)$ . Elle est définie par:

$$MZ(t_\alpha) = Z(t_\alpha) + \frac{1}{2} MSB \left( \alpha - 1 \right)^2 \quad (42)$$

avec  $MSB$  la statistique de Sargan et Bargava (1983) définie par

$$MSB = \sqrt{T^{-2} \sum_{t=1}^T y_{t-1}^2 / s^2} = MZ(t_\alpha) / MZ_\alpha \quad (43)$$

Sur le plan asymptotique, les statistiques  $MZ_\alpha$  et  $MZ(t_\alpha)$  ont les mêmes distributions asymptotiques que  $Z_\alpha$  et  $Z(t_\alpha)$  sous l'hypothèse nulle.

Les résultats des simulations réalisées par Perron et Ng (1996) montrent que les statistiques  $MZ_\alpha$  et  $MZ(t_\alpha)$  présentent à distance finie, des distorsions de niveau faibles par rapport aux statistiques  $ADF$ ,  $Z_\alpha$  et  $Z(t_\alpha)$  étudiées ci-dessus. Perron et Ng (1996) suggèrent d'utiliser une approche autorégressive pour l'estimation de la densité spectrale à la fréquence 0,  $s^2$  afin de permettre aux M-tests d'avoir de bonnes propriétés en terme de niveau. En particulier pour  $\theta = -0.8$  dans le DGP xxx ci-dessus, le niveau du test  $MZ(t_\alpha)$  est 6% pour un niveau nominal de 5% contre près de 100% pour  $MZ(t_\alpha)$  et  $Z(t_\alpha)$  lorsque  $s^2$  est estimé par la méthode de noyau.

#### 1.4. Tests d'hypothèse nulle de stationnarité KPSS

Kwiatkowski, Phillips, Schmidt et Shin (1992) proposent l'hypothèse nulle de stationnarité contre l'alternative de non stationnarité contrairement aux tests ADF et PP. Ils considèrent le modèle suivant:

$$\begin{aligned} y_t &= \beta t + x_t + u_t \\ x_t &= x_{t-1} + \varepsilon_t \end{aligned} \quad (44)$$

où  $\varepsilon_t : BB(0, \sigma_\varepsilon^2)$  et  $u_t : I(0)$ . Dans (44), le processus  $y_t$  est décomposé en trois composantes: une tendance déterministe  $\beta t$ , une marche aléatoire  $x_t$ , une erreur aléatoire stationnaire  $u_t$ .

L'hypothèse nulle de stationnarité de  $y_t$  s'écrit  $H_0 : \sigma_\varepsilon^2 = 0$  contre l'alternative de non stationnarité  $H_1 : \sigma_\varepsilon^2 > 0$ . Sous  $H_0$ , le processus  $y_t$  est  $I(0)$  si  $\beta = 0$  et  $TS$  si  $\beta \neq 0$

Le test KPSS est un test de multiplicateur de Lagrange (LM) ou un test de score donné par

$$KPSS = \left( T^{-2} \sum_{t=1}^T S_t^2 \right) / \lambda^2 \quad (45)$$

avec  $S_t = \sum_{i=1}^t u_i$  et  $\lambda^2 = 2\pi f_u(0)$  est l'estimateur de la variance de long terme de  $u_t$ .

La distribution asymptotique de *KPSS* est donnée par

$$KPSS \rightarrow \int_0^1 V^2(r) dr \quad (46)$$

où  $V(r) = W(r) - rW(1)$  si  $\beta = 0$  et  $V(r) = W(r) + (2r - 3r^2)W(1) + (-6r + 6r^2) \int_0^1 W(s) ds$  si  $\beta \neq 0$

La distribution asymptotique de KPSS est tabulée suivant les expressions de  $D(r) = 1$  ou  $D(r) = \mu + \beta t$

Comme les tests ADF, PP et leurs versions modifiées, les tests KPSS présentent aussi des problèmes de distorsion de niveau.

KPSS analyse le niveau et la puissance du test en considérant d'abord que les erreurs aléatoires  $\varepsilon_t$  sont iid. En faisant remarquer que le résidu sur lequel est fondé le test ne dépend pas de  $x_0, \beta$  et par conséquent la distribution asymptotique, ils montrent que  $\sigma_{u_t}^2$  est simplifié parce que se retrouvant au numérateur et au dénominateur. Il s'ensuit que le niveau du test ne dépend que de T et du nombre de retards utilisés pour l'estimation de  $\sigma^2 = 2\pi f_u(0)$ .

Pour ce faire, KPSS considèrent les valeurs de T=30 à 500 et trois valeurs du nombre de retards  $l$ :  $l_0 = 0, l_4 = \text{Integer}[4(T/100)^{1/4}]$  et  $l_{12} = \text{Integer}[12(T/100)^{1/4}]$  en suivant Schewert(1989).

Pour  $l_0$ , le niveau exact des tests coïncide avec le niveau nominal. Pour  $l_4$ , les 2 tests sont un peu moins précis dans les petits échantillons. le niveau exact tend lentement vers le niveau nominal lorsque T augmente. En revanche pour  $l_{12}$ , les tests KPSS présentent des distorsions importantes de niveau pour T=30 et modérées pour T=100 ou 200. Le niveau exact coïncide avec le niveau nominal pour T=500.

Lorsque l'erreur aléatoire  $\varepsilon_t$  est un *AR(1)* tel que  $\varepsilon_t = \rho \varepsilon_{t-1} + e_t$  où  $e_t : iidN(0,1)$ , KPSS considèrent les valeurs  $\rho = 0, \pm 0.2, \pm 0.5, \pm 0.8$  T entre 30 et 500. Comme attendu, le test rejette le plus souvent pour  $\rho > 0$  et rarement pour  $\rho < 0$ . Pour  $l_0$  le sur-rejet est sévère notamment pour  $\rho = 0.8$

S'agissant de la puissance des tests, KPSS considèrent les valeurs précédent de T et font varier  $\kappa = \sigma_u^2 / \sigma_\varepsilon^2$  entre 0.0001 et 10000 (10000 est pris pour l'infini). Pour T fixé, la puissance augmente avec  $\kappa$ . Elle augmente plus vite pour  $l \leq l_4$  et tend vers 1 lorsque  $\kappa$  tend aussi vers l'infini (i.e  $\sigma_\varepsilon^2 = 0$ ). Pour  $l_{12}$ , la puissance est faible dans les petits échantillons. En définitive, il y a un arbitrage à faire entre puissance et niveau du test pour le choix du nombre de retards  $l$ .

## 2. Rappel sur le test GLS

Les statistiques de test ADF et PP et leurs versions modifiées sont connues pour avoir des problèmes de niveau et de puissance. Pour régler le problème de puissance, Elliot, Rothenberg et Stock (1996) proposent un test efficace fondé sur les moindres carrés généralisés modifiant les tests de Dickey-Fuller. Les tests de ERS conduisent à des gains substantiels de puissance en présence de moyenne et de tendance inconnues.

ERS(1996) proposent les transformations suivantes pour un processus  $y_t$  de composante déterministe  $z_t = (1, t)'$ . En premier lieu, on définit  $y_t^{\bar{\alpha}}$  et  $z_t^{\bar{\alpha}}$  par

$$\begin{cases} (y_1^{\bar{\alpha}}, y_t^{\bar{\alpha}}) = (y_1, (1 - \bar{\alpha}L)y_t) & t = 2, 1, \dots, T \\ (z_1^{\bar{\alpha}}, z_t^{\bar{\alpha}}) = (z_1, (1 - \bar{\alpha}L)z_t) & t = 2, 1, \dots, T \end{cases} \quad (47)$$

avec  $\bar{\alpha} = 1 + \bar{c}/T$ . Le paramètre  $\bar{c} = -7$  si  $z_t = 1$  et  $\bar{c} = -13,5$  si  $z_t = (1, t)'$ .

En second lieu, on régresse  $y_t^{\bar{\alpha}}$  sur  $z_t^{\bar{\alpha}}$  pour obtenir  $\hat{\beta}_{GLS}$  afin de détrender  $y_t$ . On a donc

$$\hat{\beta}_{GLS} = \operatorname{argmin}_{\beta} (y_t^{\bar{\alpha}} - \beta' z_t^{\bar{\alpha}})' (y_t^{\bar{\alpha}} - \beta' z_t^{\bar{\alpha}}) \quad (48)$$

La série détrendée par GLS est donnée par

$$\hat{y}_t = y_t - \hat{\beta}_{GLS}' z_t \quad (49)$$

ERS considère comme test ponctuel réalisable, le test de rapport de vraisemblance défini par la statistique

$$P_T = [S(\bar{\alpha}) - \bar{\alpha}S(1)] / s_{AR}^2 \quad (50)$$

avec

$$S(\bar{\alpha}) = \min_{\beta} (y_t^{\bar{\alpha}} - \beta' z_t^{\bar{\alpha}})' (y_t^{\bar{\alpha}} - \beta' z_t^{\bar{\alpha}})$$

La statistique  $P_T$  n'est pas satisfaisante au vu des résultats des simulations. Elle présente de large distorsion de niveau lorsque  $k$  est sélectionné avec le critère BIC. Pour cette raison ERS suggèrent de tester l'hypothèse nulle  $H_0 : \beta_0 = \alpha - 1 = 0$  contre l'alternative  $H_1 : \alpha = \bar{\alpha}$  en considérant la régression

$$\Delta y_t = \beta_0 y_{t-1} + \sum_{j=1}^k \beta_j \Delta y_{t-j} + v_t \quad (51)$$

Les statistiques de test sont  $DF^{GLS}$  et  $MZ^{GLS}$

**Sous l'hypothèse alternative**  $\alpha = 1 + c/T$ , les statistiques  $DF^{GLS}$  et celles de la classe  $M$  ont des distributions asymptotiques dérivées dans Ng et Perron (1997). Elles sont données comme il suit:

i) Si  $z_t = 1$  pour tout  $t$

$$\begin{aligned} MZ_{\alpha}^{GLS} &\Rightarrow \left( \int_0^1 J_c(r) dW(r) \right) \left( \int_0^1 J_c^2(r) d(r) \right)^{-1} \\ MSB_{\alpha}^{GLS} &\Rightarrow \left( \int_0^1 J_c^2(r) d(r) \right)^{1/2} \\ MZ_t^{GLS} = MZ_{\alpha}^{GLS} \cdot MSB_{\alpha}^{GLS} &\Rightarrow \left( \int_0^1 J_c(r) dW(r) \right) \left( \int_0^1 J_c^2(r) d(r) \right)^{-1/2} \end{aligned} \quad (52)$$

ii) Si  $z_t = (1, t)'$  pour tout  $t$

$$\begin{aligned} MZ_\alpha^{GLS} &\Rightarrow 0.5 \left( V_{c,\bar{c}}(1)^2 - 1 \right) \left( \int_0^1 V_{c,\bar{c}}^2(r) d(r) \right)^{-1} \\ MSB_\alpha^{GLS} &\Rightarrow \left( \int_0^1 V_{c,\bar{c}}^2(r) d(r) \right)^{1/2} \\ MZ_t^{GLS} = MZ_\alpha^{GLS} \cdot MSB_\alpha^{GLS} &\Rightarrow 0.5 \left( V_{c,\bar{c}}(1)^2 - 1 \right) \left( \int_0^1 V_{c,\bar{c}}^2(r) d(r) \right)^{-1/2} \end{aligned} \quad (53)$$

où  $J_c(r)$  est le processus de Ornstein-Uhlenbeck défini par l'équation différentielle stochastique  $dJ_c(r) = cJ_c(r)dr + dW(r)$  avec  $J_c(0) = 0$  et

$$V_{c,\bar{c}}(r) = J_c(r) - r \left[ \lambda J_c(1) + 3(1-\lambda) \int_0^1 s J_c(s) ds \right], \quad \lambda = (1-\bar{c}) / (1-\bar{c} + \bar{c}^2/3)$$

Sous l'hypothèse nulle  $\alpha = 1 (c = 0)$ , les statistiques  $DF^{GLS}$  et celles de la classe  $M$  ont des distributions asymptotiques données par:

i) Si  $z_t = 1$  pour tout  $t$

$$\begin{aligned} MZ_\alpha^{GLS} &\Rightarrow \left( \int_0^1 W(r) dW(r) \right) \left( \int_0^1 W^2(r) d(r) \right)^{-1} \\ MSB_\alpha^{GLS} &\Rightarrow \left( \int_0^1 W^2(r) d(r) \right)^{1/2} \\ MZ_t^{GLS} = MZ_\alpha^{GLS} \cdot MSB_\alpha^{GLS} &\Rightarrow \left( \int_0^1 W(r) dW(r) \right) \left( \int_0^1 W^2(r) d(r) \right)^{-1/2} \end{aligned} \quad (54)$$

ii) Si  $z_t = (1, t)'$  pour tout  $t$

$$\begin{aligned} MZ_\alpha^{GLS} &\Rightarrow 0.5 \left( V_{0,\bar{c}}(1)^2 - 1 \right) \left( \int_0^1 V_{0,\bar{c}}^2(r) d(r) \right)^{-1} \\ MSB_\alpha^{GLS} &\Rightarrow \left( \int_0^1 V_{0,\bar{c}}^2(r) d(r) \right)^{1/2} \\ MZ_t^{GLS} = MZ_\alpha^{GLS} \cdot MSB_\alpha^{GLS} &\Rightarrow 0.5 \left( V_{0,\bar{c}}(1)^2 - 1 \right) \left( \int_0^1 V_{0,\bar{c}}^2(r) d(r) \right)^{-1/2} \end{aligned} \quad (55)$$

où  $V_{0,\bar{c}}(r) = W(r) - r \left[ \lambda W(1) + 3(1-\lambda) \int_0^1 s W(s) ds \right], \quad \lambda = (1-\bar{c}) / (1-\bar{c} + \bar{c}^2/3)$

**Remarque 5** Sous l'hypothèse nulle  $\alpha = 1$ , les distributions asymptotiques des statistiques de test sont les distributions ne dépend pas de  $\bar{c}$  lorsque  $z_t = 1$ .

Tableau des valeurs critiques des statistiques de test  $M^{GLS}$

Percentile	$MZ_\alpha^{GLS}$	$MZ_r^{GLS}$	$MSB^{GLS}$
	<i>Cas</i>	$z_t = 1$	
0.01	-14,327	-2,635	1,874
0.05	-8,260	-1,927	3,203
0.10	-6,024	-1,653	4,374
	<i>Cas</i>	$z_t = (1, t)'$	
0.01	-23,587	-3,414	0,144
0.05	-16,948	-2,871	0,168
0.10	-13,871	-2,591	0,185

Remarque : Les valeurs critiques des statistiques  $M^{GLS}$  ci-dessus sont obtenues par simulations de Monte Carlo en approximant le processus de Wiener  $W(r)$  par des sommes partielles de  $N(0,1)$  avec  $T = 500$  et 20000 répliques.

### Puissance asymptotique des tests $M^{GLS}$

Suivant les travaux de Ng et Perron (1998), la puissance asymptotique des tests  $M^{GLS}$  est 0,52 pour  $\bar{c} = -7.0$  correspondant à  $z_t = 1$  contre 0,36 pour les tests de la classe  $M$ . Pour  $z_t = (1, t)'$ , les puissances des tests  $M^{GLS}$  et  $M$  sont respectivement 0,48 et 0,37. Il en résulte que les tests  $M^{GLS}$  sont uniformément plus puissants que les tests de la classe  $M$  avec une fonction de puissance qui coïncide avec celle de  $DF^{GLS}$ ,

**Remarque 6** Concernant la construction de la densité spectrale à la fréquence 0 basée un

modèle AR:  $S_{AR}^2$ . Pour construire  $S_{AR}^2$ , Ng et Perron (1998) suggère d'ignorer la tendance et de maintenir la constante dans tous les cas de figure. Ils justifient ce point par le fait que le maintien de la tendance est source de plus de distorsion de niveau de test à distance finie.

La méthode de détrendisation par GLS proposée par ERS est attractive par le fait qu'elle permet d'estimer avec précision la fonction déterministe. On s'attend aussi à ce qu'elle conduise à une

réduction du biais de  $b(1)$  et par conséquent à une meilleure estimations de  $S_{AR}^2 = S_u^2 \left( 1 - b(1) \right)^2$ . L'estimateur  $S_{AR}^2$  obtenu par (0.30) est invariant à la pente de la

tendance. On pourrait aussi utiliser (51) pour estimer  $S_{AR}^2$ . Dans ce cas, Ng et Perron(1988) notent les tests  $\overline{M}^{GLS}$  au lieu de  $M^{GLS}$ .

### Problème de niveau des tests en relation avec le nombre de retards $k$

Ng et Perron(1998) ont analysé la façon dont le nombre de retards  $k$  influence les propriétés à distance finie des statistiques  $M^{GLS}$  et  $DF^{GLS}$ . Ils considèrent le processus  $y_t$  généré par

$$\begin{aligned} y_t &= d_t + u_t \\ u_t &= \alpha u_{t-1} + v_t \end{aligned} \tag{56}$$

où i)  $v_t$  est un  $MA(1)$  tel que  $v_t = e_t + \theta e_{t-1}$  avec  $e_t : iidN(0,1)$  et ii)  $v_t$  est un  $AR(1)$  tel que  $v_t = \rho v_{t-1} + e_t$ .

Pour  $T = 100$  et  $250$  et  $\theta, \rho = -0,8$  et  $0,8$ , Ng et Perron(1998) comparent pour différentes valeurs de  $k$ , le niveau exact des tests  $M^{GLS}$  et  $DF^{GLS}$  au niveau nominal de 5%. Leurs résultats montrent que pour  $\theta = -0.8$  le niveau des tests s'améliorent lorsque  $k$  et  $T$  augmentent. Par exemple, lorsque  $z_t = 1$ , le niveau exact du test  $M^{GLS}$  est 0,17 pour  $k=5$  et 0,09 pour  $k=10$  lorsque  $T=100$  contre respectivement 0,288 et 0,118 pour  $DF^{GLS}$ . Lorsque  $T=250$ , le niveau exact du test  $M^{GLS}$  est 0,188 pour  $k=5$  et 0,038 pour  $k=10$  contre 0,276 et 0,098 pour le test  $DF^{GLS}$ .

En présence de tendance, c'est à dire  $z_t = (1, t)'$ , les résultats obtenus sont similaires. Les tests  $MZ_\alpha^{GLS}$  performant mieux que les tests  $DF^{GLS}$ . Pour  $T=100$ , le niveau exact du test  $MZ_\alpha^{GLS}$  est 0,145 pour  $k=5$  et 0,141 pour  $k=10$  lorsque  $T=100$  contre respectivement 0,238 et 0,093 pour

$DF^{GLS}$ . Lorsque  $T=250$ , on a respectivement 0,096 et 0,020 pour  $MZ_{\alpha}^{GLS}$  contre 0,284 et 0,082 pour  $DF^{GLS}$ .

Les résultats présentent des tableaux différents lorsque  $v_t$  suit un  $AR(1)$  et  $\rho = 0,8$  et  $-0,80$ . Pour  $\rho = -0,8$ ,  $T = 100$  et  $z_t = 1$ , la statistique  $MZ_{\alpha}^{GLS}$  performe en niveau mieux que  $DF^{GLS}$  lorsque  $k \leq 4$ . Par exemple pour  $k = 2,3,4$  les niveaux exacts du test  $MZ_{\alpha}^{GLS}$  sont respectivement 0,033, 0,053 et 0,062 contre 0,076, 0,080 et 0,076 pour  $DF^{GLS}$ . Pour  $k \geq 5$ , le niveau exact du test  $DF^{GLS}$  varie entre 0,072 et 0,084 contre 0,074 et 0,142 pour  $MZ_{\alpha}^{GLS}$ . Lorsque  $T = 250$  le niveau exact du test  $DF^{GLS}$  correspond au niveau nominal de 5% pour tout  $k \geq 1$ . Pour  $MZ_{\alpha}^{GLS}$  le niveau exact oscille entre 0,06 et 0,096.

Lorsque  $\rho = 0,8$  et  $T = 100$ , le test  $DF^{GLS}$  performe mieux que  $MZ_{\alpha}^{GLS}$  et cette performance est d'autant meilleure que  $T$  augmente et ceci quelle que soit la valeur de  $k \geq 1$ .

### 3. Cointégration de Perron-Rodriguez 2016

Valeurs critiques asymptotiques de la statistique de test  
GLS Perron-Rodriguez 2016

m=2 (nombre de variables expliquées)	
5%	-2,354
7,5%	-3,088
10%	-2,972

Source : Perron et Rodriguez (2016)

**Annexes 4 : Tableau des indicateurs CNT actualisés**

Secteur	Conclusions du Séminaire sur les CNT, janvier 2015	Indicateur				Disponibilité
		Direct			Indirect	
		Quantité	Prix	Valeur		
<b>Agriculture</b>	Faire une enquête de la structure des coûts pour avoir une estimation de la production agricole conforme aux concepts des comptes nationaux: A : coût de production ; B : structure des coûts de production selon le type d'année (bonne, moyenne, faible) ; C : Calendrier des semis et des récoltes ou lissage.	Production agricole par spéculation // (1) Répartition infra annuelle de la production annuelle selon une grille des coûts de production; (2) Quantités produites selon le calendrier des récoltes (kg) // enquête agricole; (3) Quantités produites, collectées et/ou vendues (f/kg) // filières agriculture industrielle	(1) Prix d'offre // SIM marché agricole; (2) IHPC postes céréales non transformées, fruits et légumes	(1) Valeur de la production, des collectes et des ventes // filières agriculture industrielle	(1) Calendrier des récoltes; (2) Pluviométrie; (3) Exportations de produits agricoles (1) Population et hypothèse de productivité; (2) pluviométrie décalée pour tenir compte du calendrier des récoltes	annuel
<b>Elevage</b>	Il a été signalé aux pays utilisant le volume d'abattage comme indicateur pour estimer le cheptel de corriger le résultat par le solde du commerce extérieur de l'élevage. A : Quantité et les prix par espèces ; B : Tonnage globale ; C : Exportations, ventes de produits de l'élevage, lissage.	Effectif du cheptel // IHPI Abatage pour l'élevage (1) Abattages contrôlés (tête ou poids carcasse) (2) Sorties et entrées de cheptel (tête) (3) Production laitière	Prix moyen par espèce // (1) Prix d'offre; (2) IHPC postes viande (bœuf, mouton, chèvre et porc); volaille; (3) Prix du lait et prix des œufs			
<b>Sylviculture</b>	Il en ressort que l'un des indicateurs utilisé est la production trimestrielle du bois en grume. Le volume du bois en grume est estimé au besoin à partir du bois scié par l'utilisation d'un coefficient correcteur ; A : coût de production ; B : Utilisation de l'IPI ou des exportations de bois scié ; C : lissage.	Production forestière Production grumes				

Secteur	Conclusions du Séminaire sur les CNT, janvier 2015	Indicateur			Disponibilité
		Direct		Indirect	
		Quantité	Prix		
<b>Pêche</b>	La recommandation dégagée est d'améliorer le système de collecte des données de la pêche; les indicateurs utilisés sont les débarquements de la pêche (quantités et prix par espèce), tonnage global par mois, et le lissage ou les exportations	Production de la pêche Débarquement et/ou mise à terre (kg)		Valeur commerciale estimée	Mensuel
<b>Extraction</b>	Il ressort que les indicateurs utilisés sont l'IPPI, les quantités produites, les quantités valorisées par les prix à la production mensuelle.	IHPI Extractives Production industrielle d'or Production de pétrole brut Exportations d'or Quantités produites et/ou exportées (kg) Filières et/ou grandes entreprises	(1) IPPI	(1) Valeur de la production et/ou des exportations	
<b>Industrie</b>	Il en ressort que l'indicateur utilisé est l'IPPI à fréquence trimestrielle. Si le secteur informel est important dans la sous-branche, on utilise la tendance. Il a été recommandé de réviser l'IPPI régulièrement, de faire l'estimation de la projection à partir de la structure de la population active et de faire régulièrement des enquêtes sur le secteur informel	(1) IHPI alimentaire (agroalimentaire) (2) Production du coton retardé (2) Production du coton fibre // Production de bois scié Exportations de cacao et café transformés Importations de blé; (4) Bois transformé exporté Production de caoutchouc (2) Production d'or Production contrôlée des cuirs et peaux (unité)	(1) IPPI // INS	Calendrier des récoltes // (1) Quantité de coton égrené (kg) // GE et/ou INS; (2) Importation de produits pétroliers;	Trimestriel

Secteur	Conclusions du Séminaire sur les CNT, janvier 2015	Indicateur			Disponibilité	
		Direct		Indirect		
		Quantité	Prix			Valeur
<b>Construction</b>	Il ressort que les indicateurs utilisés sont la mise à consommation du ciment et l'IPI des matériaux de construction. Pour améliorer l'étalonnage, il a été recommandé de créer un indicateur composite de matériaux de construction et si possible inclure les coûts du travail ; une moyenne mobile des indicateurs peut améliorer les étalonnages.	IPI autre industrie manufacturière Indicateur avancé des BTP de l'IPI (1) IHPI matériaux de construction	(1) IPPI	(1) Ventes de ciment (1) Dépenses d'investissements de l'Etat	Production, importations et exportations de ciment (1) Mise à la consommation du ciment; (2) Taux d'accroissement de la population rurale	Trimestriel
<b>Electricité et Eau</b>	Il ressort que les indicateurs utilisés sont les IPI de l'eau et de l'électricité. Concernant l'électricité, il est recommandé de multiplier les quantités par les prix suivant les régimes (haute, moyenne, basse tension) avant de calculer l'indice. Pour l'eau, on utilise l'IPI directement.	(1) IHPI eau électricité (1) Consommation et importation d'électricité, consommation d'eau Quantités produites et vendues d'électricité et d'eau	(1) IPPI	(1) Valeur de la production et/ou des ventes d'électricité et d'eau // Agence de régulation et GE		Trimestriel
<b>Commerce</b>	Il ressort que les indicateurs utilisés sont les taux de marge appliqués sur la production et les importations par biens. Si le secteur informel n'est pas prépondérant, on peut essayer l'ICA déflaté de l'IHPC.		(1) IHPC dont poste entretien et réparation	(1) Chiffres d'affaires	(1) VA primaire+Secondaire (2) Taux de marge du commerce (1) Productions et imports échangées (1) Production des secteurs amonts (agriculture et industrie); (2) TVA sur consommation	
<b>Services immobiliers</b>	Il ressort que les indicateurs utilisés sont les abonnés d'électricité de moyenne et basse tension. La plupart des pays applique un lissage.	(1) IHPI énergie			(1) Taux d'accroissement de la population (1) Abonnés moyenne et basse tension (1) Stock de logements	

Secteur	Conclusions du Séminaire sur les CNT, janvier 2015	Indicateur			Disponibilité
		Direct		Indirect	
		Quantité	Prix		
<b>Activités financières</b>	Il ressort que l'indicateur utilisé est le montant des crédits à l'économie déflaté par l'IHPC. Dans une perspective où on séparerait les services bancaires et l'assurance, le lissage peut être utilisé pour traiter les assurances, et les crédits à l'économie pour les activités bancaires.	(1) Nombre des assurés		Crédits à l'économie Produits nets bancaires (1) Dépôts auprès des banques Chiffres d'affaires des assurances (3) Primes pour les assurances	Trimestriel
<b>Services de réparation</b>	il a été noté que la plupart des pays ne les dissocient pas du commerce. La présence d'un trend dans l'étalonnage de la branche commerce et réparation permet de tenir compte de ces services qui sont essentiellement informels.		(1) IHPC dont poste entretien et réparation	(1) Chiffres d'affaires  (1) VA primaire+Secondaire (2) Taux de marge du commerce (1) Productions et imports échangées (1) Production des secteurs amonts (agriculture et industrie); (2) TVA sur consommation	
<b>Hôtellerie et Restauration</b>	les indicateurs utilisés sont l'ICA déflaté pour étalonner la production, les nuitées dans les hôtels, les arrivées des touristes. Ces indicateurs peuvent être corrigés par les recettes du poste voyage de la BDP déflatées par l'IHPC.	Nuitées Arrivée de touristes Taux d'accroissement de la population Total des voyageurs	(1) Poste "Hôtel et restauration" de l'IHPC	(1) ICA hôtel // (1) Ligne "tourisme" de la Balance des paiements	(1) Indice synthétique du PIB des pays de résidence des touristes  Trimestriel
<b>Transport et entrepôts</b>	l'ICA est proposé comme indicateur, mais du fait du secteur informel prépondérant, surtout dans le transport terrestre, d'autres indicateurs comme la mise en consommation des produits pétroliers corrigés des consommations des ménages, le nombre de vignettes fiscales délivrées, le nombre de passagers et tonnage par km sont avancés.	(1) Trafics portuaires // (1) Mise à la consommation du carburant (1) Nombre de navires sortis; (2) Production transport maritime; (3) Fret aérien ou total voyageurs (1) Trafic fluvial, aérien et ferroviaire (entrées et sorties) et transit (kg) (2) Trafic routier (kg) //		(1) ICA transport  (1) VA primaire+Secondaire (1) Taxes sur transports	Mensuel, trimestriel

Secteur	Conclusions du Séminaire sur les CNT, janvier 2015	Indicateur			Disponibilité	
		Direct		Indirect		
		Quantité	Prix			Valeur
<b>Informations et Télécommunications</b>	les indicateurs souvent utilisés sont l'ICA pour étalonner la production et déflaté avec l'IHPC, les chiffres d'affaire déflatés avec l'IHPC, le nombre d'unités consommées en minutes. En définitive, l'ICA déflaté peut être utilisé pour étalonner l'agrégat annuel.	(1) Appels vers Extérieurs des Compagnies de télécoms Parc des abonnés (1) Trafic d'appels; (2) Effectifs des abonnés Durée des appels	(1) Poste "Communication" de l'IHPC	(1) Chiffre d'affaire des compagnies de télécommunication (1) ICA télécoms	(1) Taxes sur postes et communication	Trimestriel
<b>Services aux entreprises</b>	Un indicateur dérivé de la demande en CI de ces services permet d'étalonner la production de cette branche. Le Cameroun et le Burkina font simplement un lissage.	(1) IHPI énergie			(1) Taux d'accroissement de la population (1) Production des autres branches (1) Production du secteur secondaire	
<b>Services principalement non marchands</b>	Il en ressort que les indicateurs sont les effectifs des administrations publiques et les salaires déflatés avec l'IHPC. Les indicateurs sont, pour le service marchand de l'éducation, le nombre d'élèves par cycle pondérées par le coût moyen de chaque élève et étudiant. Pour le service marchand de la santé l'on peut prendre le nombre de lits occupés.	(1) Effectifs des APU	IHPC, Postes "Education", "Santé", "Loisir, spectacle et culture	Masse salariale des APU		Trimestriel
<b>Impôts et taxes</b>			IHPC global	Impôts et Taxes sur les produits Droits de douanes	(1) Productions et imports taxées	Trimestriel
<b>SIFIM</b>			IHPC	Crédits à l'économie	VA services financiers	Trimestriel