



Munich Personal RePEc Archive

## **Stability of the Structure of the Technical Coefficients: Hybrid Estimator Approach**

LATIF Laye and OUMY Laye and DIAMA Laye

University Cheikh Anta Diop of Dakar

20. June 2007

Online at <http://mpa.ub.uni-muenchen.de/3655/>  
MPRA Paper No. 3655, posted 20. June 2007

**UNIVERSITE CHEIKH ANTA DIOP DE  
DAKAR**

**STABILITE DE LA STRUCTURE DES  
COEFFICIENTS TECHNIQUES**  
Le cas du Sénégal

Dr Latif DRAMANI

Oumy LAYE

Diama Ndiaye

Juin 2007

## RESUME

Les comptes nationaux ont été, pendant longtemps, l'une des préoccupations majeures dans la production de statistiques économiques au sein de l'administration publique sénégalaise. Les travaux de la comptabilité nationale s'intéressent beaucoup à la production statistique des comptes et d'autres tableaux synthétiques comme le Tableau Economique d'Ensemble et le Tableau des Entrées Intermédiaires (TEI). Le TEI, conçu grâce à la détermination des coefficients techniques, met en exergue les interrelations entre les différentes branches d'activités.

La présente étude s'inspire des travaux de Leontief [1927] sur les modèles input-output et propose une démarche synthétique permettant de calculer des coefficients techniques robustes, ainsi que leurs intervalles de confiance. Les données utilisées proviennent d'une compilation des Tableaux Entrées Sorties (TES) de 1980 à 2004.

Pour assurer la robustesse des résultats, un estimateur hybride, combinaison optimale de la moyenne et de la médiane a été privilégié.

Les résultats obtenus mettent en évidence, une bonne adéquation entre les intervalles estimés et les données réelles observées.

Les simulations faites dans le cas d'une flambée du prix du baril en 2005, répliquent assez bien la réalité observée dans la structure économique au Sénégal.

## **TABLE DES MATIERES**

LISTE DES TABLEAUX .....	4
LISTE DES GRAPHIQUES .....	4
INTRODUCTION .....	5
<b>I-PREMIERE PARTIE</b> .....	<b>7</b>
<b>I-1 REVUE SELECTIVE DE LA LITTÉRATURE</b> .....	<b>8</b>
<b>I-2 INTERET DE L' ETUDE</b> .....	<b>12</b>
I-2-1 Les objectifs de l'étude .....	12
I-2-2 Motivations de l'étude .....	12
<b>I-3 DESCRIPTION DU TES ET DU TEI</b> .....	<b>12</b>
I-3-1 Description du TES.....	12
I-3-2 Description du TEI .....	15
I-3-3 La méthode RAS.....	16
I-3-4 Les propriétés mathématiques de la méthode RAS.....	17
<b>II-DEUXIEME PARTIE</b> .....	<b>20</b>
<b>II-1 METHODOLOGIE DE CALCUL DES COEFFICIENTS TECHNIQUES ROBUSTES</b> .....	<b>21</b>
II-1-1 Le schéma synthétique de la ventilation.....	21
II-1-2 Détermination de la structure des $\beta_i$ .....	22
II-1-3 Détermination de la structure des $\gamma_{\bullet j}$ .....	22
II-1-4 Le processus d'élaboration des coefficients techniques robustes.....	23
<b>II-2 LA DETERMINATION DE L'INTERVALLE DE CONFIANCE</b> .....	<b>26</b>
II-2-1 Construction d'un intervalle de confiance à partir d'un pivot.....	26
II-2-2 Construction d'un intervalle de confiance empirique.....	28
<b>TROISIEME PARTIE</b> .....	<b>30</b>
<b>III-1 LES PERFORMANCES DU MODELE</b> .....	<b>31</b>
<b>III-2 ESTIMATION</b> .....	<b>32</b>
III-2-1 Estimation du TES robuste.....	32
III-2-2 Commentaire des résultats du tableau 7 ( voir annexe).....	32
<b>III-3 SIMULATIONS ET COMPARAISONS</b> .....	<b>34</b>

III-3-1 Choix de l'estimateur.....	34
III-3-2 Simulation d'une augmentation de 10% du prix du baril du pétrole.....	35
<b>LIMITES DE L'ETUDE ET CONCLUSION.....</b>	<b>37</b>
<b>BIBLIOGRAPHIE.....</b>	<b>39</b>
<b>ANNEXES .....</b>	<b>41</b>
ENCADRE1 : DISTRIBUTION DU RAPPORT DE DEUX VARIABLES .....	42
ENCADRE 2 : LES JUSTIFICATIFS DU CHOIX DE L'ESTIMATEUR.....	44
ENCADRE 3 : TEST DE NORMALITE.....	46
ENCADRE 4 : TEST DE STATIONNARITE .....	47
ENCADRE 5 : L'ESTIMATEUR COMBINAISON LINEAIRE DE LA MOYENNE ET DE LA MEDIANE .....	49

#### **LISTE DES TABLEAUX**

Tableau 1: Test d'adéquation et de performance des estimations sur les deux estimateurs retenus .....	34
Tableau 2 : Simulation d'une augmentation du prix de l'énergie.....	35
Tableau 3 : Test de normalité sur les branches .....	51
Tableau 4 : Liste des sigles.....	52
Tableau 5 : Calcul des erreurs quadratiques moyennes (RMSE) des produits disponibles par branches.....	53
Tableau 6 : Calcul des erreurs quadratiques moyennes des consommations intermédiaires par branches .....	54
Tableau 7 : Coefficients Techniques de 2001 à 2004 et intervalles de confiance.....	55
Tableau 8 : Structure médiane de la part des produits dans la consommation intermédiaire .....	56

#### **LISTE DES GRAPHIQUES**

Graphique 1 : description schématique du TES .....	13
Graphique 2 : Tableau entrées sorties .....	14
Graphique 3 : Illustration du Calcul du TEI.....	19
Graphique 4 : Schéma de la ventilation.....	23
Graphique 5 : Courbe d'une loi normale et illustration des bornes de l'intervalle de confiance .....	27
Graphique 6 : Illustration des lois suivis par les deux estimateurs .....	29

## INTRODUCTION

Les comptes nationaux ont été, pendant longtemps, l'une des préoccupations majeures dans la production de statistiques économiques au sein de l'Administration Publique. Ils donnent une vision globale de la situation économique du pays en retraçant l'activité des principaux agents (en particulier, les entreprises financières ou non financières, les administrations publiques et privées, les ménages). Ils sont un outil essentiel d'aide à la prise de décision à la disposition des autorités politiques.

Les comptes nationaux fournissent un ensemble de tableaux synthétiques : le Tableau des Ressources et des Emplois (TRE), le Tableau Entrées Sorties (TES), le Tableau des Comptes Economiques Intégrés (TEI), la Matrice de Comptabilité Sociale (MCS) etc.

Le Tableau Entrées et Sorties est composé du Tableau des Entrées Intermédiaires (TEI) et de la matrice des emplois finals. Le TEI, conçu grâce à la détermination des coefficients techniques, met en exergue les interrelations entre les différentes branches d'activités. En ligne, il fournit la ventilation et la consommation intermédiaire totale d'un produit par les différentes branches utilisatrice. En colonne, le TEI donne la décomposition de la consommation intermédiaire totale d'une branche selon les différents produits.

Toutefois, en dépit des efforts faits, le Bureau de la comptabilité nationale ne disposait pas de l'ensemble des tableaux synthétiques recommandés par le SCN93. Cependant, des efforts ont été réalisés dans un passé récent, pour produire annuellement des TES et des TRE.

Par ailleurs, le retard dans la production des comptes économiques a été rattrapé. Ainsi, les comptes provisoires, semi définitifs et définitifs sont produits en respectant les délais internationaux. Des TES ont été choisis pour servir de base de calcul aux coefficients techniques pour l'élaboration des TES en valeur et en volume.

Cette étude a pour objectif principal de tester la stabilité des coefficients techniques et d'estimer un TEI robuste, c'est-à-dire un TEI dont les coefficients techniques sont stables dans le temps. Les objectifs spécifiques sont d'abord le calcul des coefficients qui permettront d'évaluer le total des consommations intermédiaires par branche d'activités, ensuite la détermination des consommations intermédiaires par produit et par branche. Enfin, une simulation d'un TEI sera faite après avoir confronté sur une période récente les paramètres (coefficients techniques) à la réalité.

La robustesse des coefficients techniques choisis n'a jamais été démontrée et leur non pertinence peut entraîner des erreurs dans l'estimation des différentes grandeurs macro-économiques. C'est pourquoi, des travaux sur la robustesse et l'estimation de ceux-ci seront utiles pour améliorer la qualité des données des comptes nationaux.

Les coefficients techniques, dont les intervalles de variation sont déterminés dans ce document, sont les paramètres qui permettent d'avoir une bonne ventilation des consommations intermédiaires au sein des branches de l'activité productive.

Cette étude est présentée en trois parties.

La première partie fait une description du TES et du TEI après avoir présenté une revue succincte et sélective de la littérature sur l'analyse input output.

La deuxième partie décrit la méthodologie de conception du TES robuste, ainsi que la spécification des intervalles de confiance selon deux approches différentes.

- la première fondée sur une fonction pivotale
- la seconde fondée sur un mélange de lois

La troisième partie examine les résultats des estimations, et propose des simulations sur le TES généré. Une série de tests statistiques est effectuée pour confronter les estimations obtenues aux réalités empiriques.

## **I-PREMIERE PARTIE**



## I-1 REVUE SELECTIVE DE LA LITTÉRATURE

Si l'on trouve dans les articles de Leontief datant des années trente (Leontief [1936] [1937] [1941]), l'essentiel du matériau analytique qui fait l'analyse input-output \_ la combinaison d'un tableau de comptabilité nationale et d'un modèle mathématique utilisant le calcul matriciel \_ l'analyse input-output en tant que telle est le résultat de la collaboration, à la demande du gouvernement américain puis de l'armée de l'air américaine, entre Leontief et le Bureau of Labor Statistics, de 1941 à 1954.

L'approche input-output permet de construire dans un premier temps un tableau des échanges interindustriels où figurent non seulement l'*output* de chaque industrie mais aussi la quantité de bien des autres industries que chaque industrie consomme. On peut alors calculer la quantité de bien d'une autre industrie ( $j$ ) nécessaire à la production d'une unité de bien de chaque industrie ( $i$ ) : ce ratio est un coefficient technique<sup>1</sup>. Pour Leontief, les coefficients techniques sont directement tirés des Eléments d'économie politique pure de Walras<sup>2</sup>.

D'ailleurs, l'analyse de la production de Walras est formellement similaire au modèle input-output de Leontief<sup>3</sup>.

Le modèle de Leontief part de l'écriture des égalités emplois-ressources pour chaque industrie d'une part et des égalités entre les valeurs ajoutées additionnées aux coûts de production et les prix d'autres part, formant deux systèmes d'équations autonomes. En faisant apparaître les coefficients techniques dans le système des quantités, on obtient l'écriture matricielle suivante : en notant  $\mathbf{d}$  le vecteur colonne des demandes finales par industrie,  $\mathbf{A}$  la matrice des coefficients techniques (la matrice structurelle),  $\mathbf{X}$  le vecteur des *outputs*,  $\mathbf{P}$  le vecteur-prix,  $\mathbf{R}$  le vecteur des valeurs ajoutées (salaires et profit) et enfin,  $\mathbf{I}$  la matrice identité, on a :

$$(\mathbf{I}-\mathbf{A})\mathbf{X} = \mathbf{d} \text{ et } (\mathbf{I}-\mathbf{A})'\mathbf{P} = \mathbf{R}$$

C'est ce modèle et le tableau entrées-sorties associé qui prennent forme aux Etats-Unis dans les années trente et quarante. Leontief indique que l'idée de l'analyse input-output était déjà formée à son arrivée aux Etats-Unis en 1931, lors de son travail au centre de recherche sur l'économie mondiale de Kiel, à la fin des années vingt. Pour ce dernier, c'est son travail sur l'estimation économétrique des fonctions d'offre et de demande en équilibre partiel qui l'aurait amené à envisager la supériorité de l'équilibre général et à en proposer une version simplifiée et opérationnelle avec l'analyse input-output<sup>4</sup>.

---

<sup>1</sup> Leontief fait l'hypothèse que les coefficients techniques sont constants, c'est-à-dire qu'il n'y a pas de substitution entre les facteurs de production et que les rendements sont constants. Par ailleurs, Leontief suppose qu'il n'y pas de productions jointes. Dans le modèle ouvert, le travail est le seul facteur non produit (facteur primaire).

<sup>2</sup> On trouve aussi dans des publications précoces la référence à Cassel (Leontief [1927]).

<sup>3</sup> Le tableau entrées-sorties associé au modèle mathématique est appelé analyse *input-output*. Le tableau seul ne constitue pas l'analyse *input-output*. Ce n'est d'ailleurs pas le tableau seul qui a fait le succès de l'analyse *input-output* mais bien la combinaison du tableau et du modèle. Sans le modèle on ne peut effectuer ni prévision, ni planification ou mettre en oeuvre des politiques économiques.

<sup>4</sup> Ce n'est que très tardivement que Leontief a reconnu les racines classiques de son approche. On peut penser que Leontief a emprunté les formalismes et certaines analyses de Walras pour exprimer des idées plutôt « classiques ». D'où la troublante ressemblance entre le modèle de Leontief et les (tardives) équations de Sraffa dans *La production de marchandises par des marchandises*. Ce

Pour Leontief c'est en 1928- 1929 que se forme l'idée de l'analyse input-output alors qu'il travaille sur l'estimation des fonctions d'offre et de demande à l'institut d'économie mondiale de Kiel. Par ailleurs Leontief affirme dès le début des années trente que l'analyse input-output est une application de la théorie de l'équilibre général de Walras et reprend le Tableau économique de Quesnay (1758).

En effet, quatre interprétations différentes se forment quant à l'origine de la formation de l'analyse input-output:

- les économistes soviétiques ont inventé l'analyse input-output dans les années vingt ;
- c'est dans le cadre de la préparation du plan quinquennal sous Staline, à la fin des années vingt, qu'a été inventé l'analyse input-output;
- c'est Leontief qui, à partir de ces travaux et dès les années vingt, a inventé l'analyse input-output ;
- il n'y a aucune influence des travaux des économistes soviétiques.

Ces différentes opinions ont été adoptées par divers économistes depuis les années soixante jusqu'à l'effondrement de l'URSS (Levine [1962] et [1964]) (Jasny [1962]) (Lange [1969]) (Spulber et Dadkhah [1975]) (Tretyakova et Birman [1976]) (Clark [1984]) (Belykh [1989]).

Herbert S. Levine suit Leontief dans son interprétation du rôle joué par les travaux des économistes soviétiques dans ses propres recherches sur l'analyse input-output. Levine comme Leontief considère comme négligeable sinon nulle la part due aux économistes soviétiques des années vingt dans l'invention de l'analyse input-output.

L'un comme l'autre insiste plutôt sur le rôle du tableau économique de Quesnay (pour le tableau entrées-sorties) et celui de la théorie de l'équilibre général de Walras (pour le modèle input-output de Leontief) :

*« S'il fallait choisir une influence qui puisse être considérée comme dominante, le choix à partir des faits disponibles serait Walras et son approche mathématique par l'équilibre général de l'analyse économique »*

Au delà des enjeux factuels évidents, la question posée dépasse celle de la paternité d'une idée. La puissance de l'analyse input output s'est développée dans les périodes de crises, et a permis par ricochet de développer des disciplines comme la comptabilité nationale.

La construction de systèmes de comptabilité nationale dans l'entre-deux-guerres vise ainsi, à donner une image chiffrée de la structure économique et des flux qui l'animent. Si la macroéconomie se doit de proposer une analyse des cycles, la comptabilité nationale doit permettre de procéder à la mesure effective des désajustements engendrés par les cycles. Cette ambition empirique oriente alors les recherches en comptabilité nationale, de Mitchell [1927, 1946] ou de Maurice Copeland[1954] aux Etats-Unis, d'Arthur Bowley[1924, 1927, 1939 ] et de Colin Clark[1924-31-32-37-39] en Grande-Bretagne, d'André Vincent en France, d'Edvan Cleeff[ ] aux Pays-Bas, d'Erik Lindahl[1957 ] en Suède ou de Frisch en Norvège. Les auteurs comme Odd Aukrust et Petter Jakob Bjerve[ ] orientent résolument les comptes nationaux vers une mesure des grands

---

n'est pas l'examen des origines soviétiques qui permet de lier le modèle de Leontief à la théorie classique mais les travaux berlinois de Leontief [1927] [1928].

agrégats économiques, qui est inscrite dans une compréhension des flux marchands et financiers. La comptabilité nationale s'impose peu à peu non seulement comme le pendant empirique de la théorie macroéconomique, mais aussi comme l'un des instruments indispensables de la politique économique.

La problématique de la mise en place des tableaux entrées-sorties est inspirée des travaux de planification de l'économie soviétique.

Les origines profondes de l'analyse input output remontent aux relations entre les travaux qui ont conduit à la naissance de la programmation linéaire, en Union Soviétique et aux Etats-Unis à la même période (influence, ignorance, convergence). Il convient à cette occasion, de faire retour sur une possible origine soviétique de l'analyse input-output dont Leontief aurait pu trouver l'inspiration dans les conceptions soviétiques des années 1920 telles qu'elles se sont exprimées dans le premier bilan de l'économie nationale de l'URSS pour les années 1923- 1924.

Quoique initialement conçue dans des buts concrets, l'analyse input-output fait apparaître des liens étroits entre le développement des mathématiques et celui de l'analyse économique la plus abstraite puisqu'elle inspirera les démonstrations modernes d'existence de l'équilibre général des années 1950. **Leontief (1986)** explique que les avancées de l'analyse input-output se sont faites au rythme des financements et des élections politiques aux Etats-Unis. L'analyse input-output est en effet coûteuse en moyens financiers, humains et techniques. Le problème de la résolution matricielle du modèle a nécessité la construction de modèles très agrégés (10 secteurs), donc faiblement opératoires, jusqu'à ce que des ordinateurs soient construits et que soient développés des programmes informatiques capables d'intégrer des dizaines puis des centaines de branches dans les tableaux input-output.

Si le modèle de Leontief entame, au milieu des années 1940, une nouvelle orientation, dans la recherche fondamentale, cela est lié aux travaux de Dantzig lorsqu'il travaillait, sous la direction de Leontief, pour l'armée de l'air américaine. En 1947, Dantzig rencontre Koopmans qui devient l'année suivante directeur de la Cowles Commission. C'est là qu'ils pourraient avoir découvert leur intérêt commun pour les modèles de planification. En 1949, Dantzig publie deux articles sur la programmation d'activités interdépendantes où le « modèle mathématique présenté est une généralisation du modèle de Leontief. Il est étroitement lié à celui que Von Neumann a présenté dans l'article intitulé "A Model of General Equilibrium" (1935-36) » [**Dantzig-Wood, 1949**]. Koopmans va utiliser cette nouvelle problématique et la voie ouverte par Dantzig : ce sont les débuts de l'analyse d'activité. Le point focal de cette période est le séminaire de la Cowles Commission et de la RAND. Participent à ce séminaire Dantzig (simplexe, modèle de Leontief généralisé), ainsi que Arrow, Koopmans et Samuelson (théorème de non-substitution). Autour de la Cowles Commission, Debreu et McKenzie travaillent également sur des modèles de Leontief. C'est dans ce cadre que seront mises au point des avancées techniques mathématiques concernant les espaces convexes, l'optimisation, le dual ainsi que des

améliorations de la méthode du simplexe. La portée de ces avancées aboutira finalement aux démonstrations d'existence par Arrow et Debreu et par McKenzie en 1954.

Ces développements, avant tout mathématiques, répondent aux désirs de mathématiciens « effrayés » par la faiblesse de la formalisation des modèles économiques. Ils coïncident aussi avec la polémique lancée par Koopmans [1947] sur la « mesure sans théorie » contre le NBER (National Bureau of Economic and Research). La position de Leontief sur la méthode à adopter pour l'économie reste assez différente: l'économie théorique ne peut avancer sans l'économie appliquée et aux concepts doivent correspondre des entités construites mesurables [Leontief, 1958]. Leontief critique sévèrement l'insistance excessive de Koopmans sur la rigueur logique des modèles et la séparation entre théoriciens et chercheurs empiriques. L'urgence, dit Leontief, est à la recherche de données empiriques nouvelles et non aux sophistications formelles. Pourtant, les chercheurs de la Cowles Commission ne retiennent de l'analyse input-output que le modèle (théorique) de Leontief. En même temps, se forme un nouveau champ disciplinaire, la programmation linéaire et émerge l'idée que la programmation linéaire, l'analyse input-output et la théorie des jeux relèvent d'une même matrice intellectuelle (Dorfman, Samuelson et Solow, 1958 et J.R Hicks, 1960).

C'est aussi la richesse de l'analyse input-output que de pouvoir ramener aux questions les plus abstraites posées par l'équilibre général : les démonstrations d'existence ou les débats épistémologiques sur le rapport des énoncés théoriques avec les phénomènes.

La construction des premiers tableaux entrées-sorties s'est faite indépendamment des premières ébauches de comptabilité nationale. Sur ce point les expériences nationales ont été diverses. Dans les pays scandinaves, sous l'impulsion de Frisch, la comptabilité nationale a été construite pour répondre aux exigences de l'économétrie naissante, en référence à une macroéconomie directement inspirée de l'équilibre général. En France, où l'analyse input-output bénéficie, juste après la seconde guerre mondiale, de l'engouement pour la planification, l'INSEE intègre celle-ci à la comptabilité nationale d'abord sous la forme d'un TEI puis d'un TES. L'intégration de l'analyse input-output dans les systèmes nationaux de comptabilité et sa généralisation internationale (ONU, OCDE), au début des années soixante, a inauguré un nouveau corpus de connaissances. Aujourd'hui, la plupart des modèles économiques intègrent des TES à des niveaux de détail plus ou moins fins. L'usage de TES comme outil de planification et de prévision est intrinsèque à l'analyse input-output. Ceci tient à sa combinaison de modélisation et d'organisation logique des données. L'intégration des TES ne se limite cependant pas aux modèles de Leontief et, à partir des années soixante, on voit se diversifier les modèles ayant recours au TES. L'un des premiers projets de modélisation inspiré du TES est la construction, à l'initiative de Stone, des Social Accounting Matrix (SAM); les SAM ont l'avantage de distinguer différents groupes d'agents de manière à évaluer les transferts des uns vers les autres. Cette approche connaît un vif succès auprès d'institutions telles que la Banque Mondiale et l'ONU. Les SAM, en orientant l'analyse *input-*

*output* vers la thématique de la croissance, portent un projet plus large d'intégration des données de la comptabilité nationale que l'on retrouvera dans les modèles d'équilibre général calculables des années 1970.

## **I-2 INTERET DE L' ETUDE**

### **I-2-1 Les objectifs de l'étude**

Les travaux de la comptabilité nationale s'intéressent à la production statistique des comptes et d'autres tableaux synthétiques comme les TEE, les TRE et les TEI. La présente contribution s'écarte de cette routine et propose une démarche synthétique permettant de calculer des coefficients techniques robustes.

L'objectif principal de cette étude est de proposer une méthodologie de calcul des coefficients techniques robustes. Il s'agira spécifiquement de :

- choisir un estimateur robuste pour l'estimation des coefficients techniques ;
- effectuer une simulation sur les données réelles du TES de 2001 à 2004 ;
- produire une estimation des coefficients techniques robustes ;

Les données utilisées proviennent d'une compilation des TES de 1980 à 2004, soit une série de 25 années de TES.

### **I-2-2 Motivations de l'étude**

La question centrale qui se pose est la détermination et la quantification des différentes interrelations qui existent dans le TEI.

La présente étude constitue un pan important dans la mise en oeuvre des comptes trimestriels. Elle a été élaborée, afin de mettre à la disposition du Bureau de la comptabilité nationale, les structures de coefficients techniques et de ventilation permettant de générer rapidement des comptes trimestriels au Sénégal.

La méthodologie proposée permet d'éviter le travail fastidieux qui consisterait à estimer les 1681 séries chronologiques de coefficients techniques, provenant du TES. Ainsi, l'étude est faite sur 41 séries regroupant les totaux des consommations intermédiaires par branche, ce qui permet de raccourcir les délais dans les estimations.

## **I-3 DESCRIPTION DU TES ET DU TEI**

### **I-3-1 Description du TES**

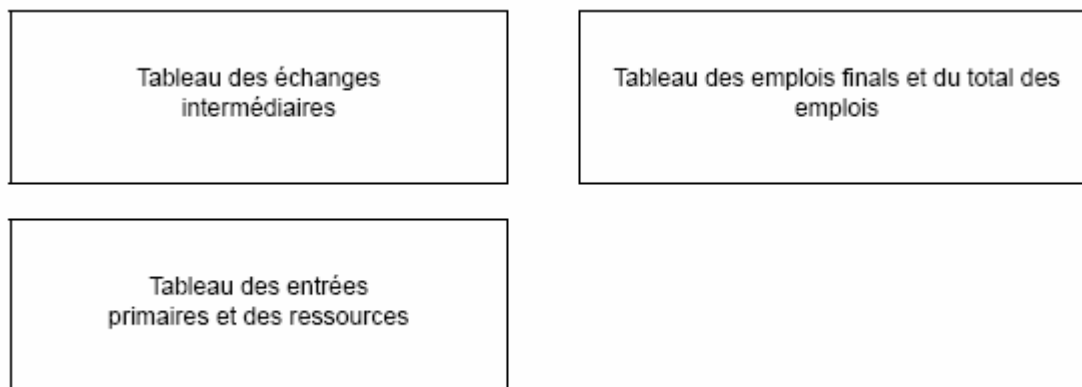
Le tableau entrées-sorties est un tableau qui fournit une description des activités de production intérieures et des opérations sur produits d'une économie. Il permet de voir l'interdépendance existant entre les différentes branches d'activité d'une économie en indiquant les biens et services qui sont entrés dans la production d'autres biens et services et les utilisations finales des produits. Il donne également la structure des coûts de production.

Le tableau entrées-sorties symétrique branche x branche<sup>1</sup> ou produit x produit repose sur le regroupement d'unités de production homogène en branches homogènes. Les unités de production homogène sont des unités qui n'exercent qu'une seule activité, les branches d'activité s'identifiant ainsi aux produits.

Le tableau entrées-sorties se décompose en trois sous-tableaux: le tableau des échanges intermédiaires, le tableau des emplois finals et le tableau des entrées primaires et des ressources<sup>3</sup>.

Ces trois sous-tableaux s'articulent de la façon suivante:

**Graphique 1 : description schématique du TES**



**Le tableau des échanges intermédiaires (TEI):** ce tableau contient l'ensemble des biens et services qui entrent dans le processus de production d'autres biens et services.

Il constitue la partie centrale du tableau entrées-sorties. Il s'agit d'un tableau carré comprenant autant de lignes que de colonnes. Chaque ligne ou colonne correspond à une branche d'activité ou à un produit. Une ligne décrit les livraisons intermédiaires d'un produit donné à toutes les branches de l'économie.

Une colonne représente les emplois des différents produits nécessaires à la production d'une branche donnée.

**Le tableau des emplois finals:** ce tableau contient les emplois finals par produit. Il présente en lignes les produits et en colonnes les catégories de la demande finale. Il répartit les livraisons à la demande finale entre la consommation privée ou publique, la formation brute de capital fixe, la variation des stocks et les exportations. Une colonne supplémentaire reprend le total des emplois (intermédiaires et finals) par produit.

**Le tableau des entrées primaires et ressources:** ce tableau présente, en colonnes, les branches d'activité ou produits et, en lignes, les composantes de la valeur ajoutée (rémunérations des salariés, consommation de capital fixe, impôts nets liés à la production, excédent net d'exploitation) et les transferts<sup>1</sup> et importations.

Une dernière ligne reprend le total des ressources. Les colonnes donnent pour chaque produit, la ventilation de la valeur ajoutée en ses différentes composantes, plus les transferts et les importations de produits similaires. Les lignes présentent la répartition de chaque composante de la valeur ajoutée entre les différentes branches de l'économie.

Le tableau entrées-sorties contient l'ensemble des flux qui constituent le compte de biens et services, le compte de production et le compte d'exploitation.

**Le compte de biens et services (C0)** met en regard, pour l'ensemble de l'économie et par branche, les ressources et les emplois de biens et services.

**Le compte de production (C1)** retrace les opérations qui constituent le processus de production proprement dit. Le côté "ressources" enregistre la production et le côté "emplois", la consommation intermédiaire. Son solde est la valeur ajoutée.

**Le compte d'exploitation (C2)** enregistre les opérations de répartition qui sont directement liées au processus de production. Il reprend en ressources la valeur ajoutée et les subventions d'exploitation et en emplois les impôts liés à la production et la rémunération des salariés. Son solde est l'excédent brut d'exploitation.

Le tableau entrées-sorties se représente sous une forme simplifiée de la manière suivante:

**Graphique 2 : Tableau entrées sorties**

	Produits	Emplois finals	Total des emplois
Produits	Consommation intermédiaire (CI)	Consommation finale publique et privée (C) Formation brute de capital (FBCF) Variation des stocks ( $\Delta S$ ) Exportations (EXP)	Total des emplois par produit
Valeur ajoutée	Composantes de la valeur ajoutée (VA)		
$\Sigma$	Production effective (Pe)		
Transferts	Transferts		
$\Sigma$	Production distribuée (P)		
Importations	Importations (IMP)		
$\Sigma$	Total des ressources par produit		

Comme le montre le tableau entrées-sorties ci-dessus, une ligne représente les sorties ou emplois de la branche. Le total des emplois d'un produit est égal à la somme des livraisons de ce produit à toutes les branches utilisatrices et aux composantes de la demande finale. En colonne apparaissent les entrées de la branche.

Elles se composent des achats de produits intermédiaires<sup>1</sup> aux autres branches et des entrées primaires (les différentes composantes de la valeur ajoutée). La somme de ces deux éléments est égale à la production effective de la branche ( $CI + VA = Pe$ ). Pour obtenir le total des ressources d'un produit, il faut y ajouter les transferts et les importations de produits similaires.

Pour le produit de chaque branche, le total des ressources est égal au total des emplois:

$$P + IMP = CI + C + FBCF + \Delta S + EXP$$

Pour les besoins de l'analyse économique, le tableau entrées-sorties tel que décrit plus haut doit être complété par *deux tableaux supplémentaires*: un tableau des emplois des importations et un tableau entrées-sorties pour la production intérieure.

Le tableau des emplois des importations détaille, par produit, les utilisations intermédiaires et finales des importations. Il se compose des sous-tableaux des échanges intermédiaires et des emplois finals. Le tableau entrées-sorties pour la production intérieure détaille les emplois des produits nationaux. Il se présente sous la même forme que le tableau entrées-sorties total.

Lors de la réalisation d'un tableau entrées-sorties se pose le problème du choix du système d'évaluation des flux. Le tableau entrées-sorties 1985 a été réalisé aux prix départ-usine hors TVA pour éliminer l'influence de ce type de fiscalité indirecte sur le processus de production. Le prix départ-usine correspond au prix de vente du producteur, augmenté des impôts nets des subventions acquittés par ce producteur. Il ne comprend pas les coûts de distribution représentés par les frais de transport et les marges de commerce.

### **I-3-2 Description du TEI**

Le Tableau des Echanges Intermédiaires (TEI) est la pièce centrale du TES. Il représente les livraisons des branches aux autres branches et c'est lui qui finalement réconcilie l'offre et la demande.

Son élaboration à partir de données complètes aurait exigé beaucoup de temps et de ressources; c'est la raison pour laquelle nous avons opté pour une technique de mise à jour moins exigeante puisqu'elle n'impose de connaître que les totaux marginaux de la nouvelle matrice à estimer.

Dans cette perspective, les totaux en colonnes (entrées intermédiaires) ont déjà pu être évalués dans les comptes de production, tandis que les totaux en lignes (sorties intermédiaires) ont pu faire l'objet d'une première évaluation grâce aux données disponibles.

L'équilibrage final n'a cependant pas encore été réalisé.

La présente section sera consacrée à décrire la façon dont cet équilibrage a été réalisé. Il montrera ensuite comment ces résultats, combinés au TEI de 1980 ont pu être exploités pour estimer, grâce à la méthode RAS, les cases individuelles de la matrice des échanges intermédiaires du TES robuste.

Le tableau d'absorption détaille les emplois intermédiaires de chaque produit (biens et services), d'origine nationale ou importée par les branches d'activité. Si on y ajoute la demande finale par produit, on obtient l'emploi total par produit. La différence entre un tableau de formation et d'absorption et un tableau entrées-sorties symétrique (produit x produit) est que les branches d'activité ne sont pas nécessairement homogènes. Cela signifie qu'à côté de l'activité principale qui les caractérise, les entreprises peuvent avoir des productions secondaires qui sont le produit caractéristique d'autres branches d'activité, et c'est ce que nous apprend la matrice de formation.

Cette hétérogénéité s'explique par le fait que les tableaux de formation et d'absorption sont établis sur base d'unités statistiques observables, telles que les entreprises, les établissements, qui, le plus souvent, ne sont pas



homogènes. Les branches homogènes constituent une construction artificielle. Elles regroupent des unités de production homogène.

L'homogénéisation complète des matrices des échanges intermédiaires et des entrées primaires repose sur des données supplémentaires ou sur des hypothèses relatives à la technologie. Les deux hypothèses classiques sont les suivantes:

- l'hypothèse d'une **technologie unique par produit**: les biens et services intermédiaires et les entrées primaires qui entrent dans le processus de fabrication d'un produit sont les mêmes quelle que soit la branche qui les produit.
- l'hypothèse d'une **technologie unique par branche d'activité**: une branche d'activité utilise la même combinaison d'entrées pour tous ses produits.

Pour procéder à l'homogénéisation complète d'un tableau entrées-sorties, c'est le principe d'une technologie unique par branche d'activité qu'il convient de retenir. En effet, en général près de 75% de la production secondaire consiste en des produits fatals (gaz de mine, de cokeries, de haut-fourneaux,...) et des produits voisins qui, par définition, sont le fruit de la même technologie que celle utilisée pour la fabrication du produit principal. Il semble donc logique d'adopter le principe d'une technologie unique par branche d'activité plutôt qu'une technologie unique par produit.

On peut se poser la question de l'intérêt de cette méthode d'homogénéisation d'un tableau entrées-sorties. Nous avons décidé d'y recourir pour les raisons suivantes:

- en négligeant ce problème, on fait consciemment une faute méthodologique;
- un tableau de formation-absorption ou un système entrées-sorties n'a pas de solution. Même si le nombre de produits est égal au nombre de branches d'activité, on a un système de  $n$  équations à  $2n$  inconnues: la production de  $n$  produits et la production de  $n$  branches d'activité;
- les transferts concernent principalement quelques branches d'activité. Pour celles-ci, l'erreur n'est pas négligeable;
- lorsque l'on utilisera les tableaux entrées-sorties comme instrument d'analyse, on calculera des effets multiplicateurs. Une erreur commise au niveau du tableau entrées-sorties lui-même s'en trouverait dès lors multipliée: elle augmenterait non seulement en ampleur mais elle risquerait aussi d'affecter d'autres éléments qui étaient correctement estimés dans la matrice de départ.

### **I-3-3 La méthode RAS**

Au départ, on ne connaît de la matrice des échanges intermédiaires de l'année  $K$  que les totaux en lignes et en colonnes. Il s'agit respectivement des livraisons et des consommations intermédiaires par branche. On ne connaît pas les éléments individuels qui composent la matrice mais il est possible de les estimer si l'on utilise l'information qui nous est fournie par la matrice des échanges intermédiaires d'une année antérieure  $K_0$ .

Il existe différentes méthodes qui permettent d'estimer une matrice dont on connaît les totaux marginaux, sur base d'une estimation initiale qui ne répond pas aux conditions marginales. Ces méthodes ont en commun la recherche d'une matrice qui diffère le moins possible de l'estimation initiale tout en satisfaisant aux conditions marginales.

La méthode RAS est la méthode la plus fréquemment utilisée car elle possède des propriétés mathématiques fort intéressantes et elle est aisée à appliquer. Elle est principalement associée à la mise à jour des tableaux entrées-sorties mais elle peut être appliquée à d'autres problèmes se présentant sous forme matricielle (matrices qui ne doivent pas obligatoirement être carrées).

La méthode est un processus itératif qui consiste à opérer un ajustement proportionnel sur les lignes et sur les colonnes de la matrice de base, jusqu'à concordance.

Lorsque l'on utilise la méthode RAS pour l'actualisation des tableaux entrées-sorties, l'estimation initiale qui sert de point de départ à la technique de mise à jour est basée sur les données de l'année de base, ici l'année  $K_0$ .

La méthode RAS "modifiée" permet l'adjonction de données exogènes à la méthode RAS initiale. Si certains éléments de la matrice des échanges intermédiaires peuvent s'obtenir de façon exogène, il serait en effet malheureux de ne pas en tenir compte. Concrètement, on place des zéros dans les cases de la matrice initiale pour lesquelles on dispose d'informations assurées. On corrige les totaux marginaux en fonction et on applique ensuite la méthode RAS sur la matrice modifiée.

On replace enfin les données exogènes dans les cases appropriées. L'adjonction de données exogènes à la méthode RAS permet d'augmenter la qualité de l'actualisation.

### **I-3-4 Les propriétés mathématiques de la méthode RAS**

Le problème de départ peut être posé de la manière suivante: on minimise la distance entre l'estimation, représentée par la matrice  $X^*$ , et la matrice initiale  $X_0$  sous la contrainte que l'estimation satisfasse aux conditions marginales de la matrice inconnue  $X$ .

$$\text{Min } d(X^*, X_0)$$

$$X^* \cdot i = X \cdot i, \quad i' \cdot X^* = i' \cdot X$$

où  $i$  est un vecteur unité dont la dimension est égale au nombre de lignes ou de colonnes de la matrice par laquelle il est multiplié.

La méthode RAS offre une solution  $X^*$  qui est bi-proportionnelle à l'estimation initiale  $X_0$ :

$$X^* = \hat{r} \cdot X_0 \cdot \hat{s}$$

En pré et en post-multipliant la matrice initiale respectivement par les matrices diagonales des multiplicateurs de lignes ( $r$ ) et des multiplicateurs de colonnes ( $s$ ), on obtient une estimation  $X^*$  de la matrice inconnue  $X$ , qui satisfait aux conditions marginales. C'est pour cette raison que la méthode RAS est qualifiée de méthode bi-proportionnelle.

Si l'on pose comme condition que les éléments des matrices  $X_0$  et  $X^*$  sont non négatifs (plus grands ou égaux à zéro), la méthode RAS donne une solution unique au problème (il n'existe qu'une matrice non négative  $X^*$  qui satisfasse aux conditions marginales et qui soit bi-proportionnelle à la matrice initiale  $X_0$ ).

Etant donné que le système est non-linéaire, les valeurs des vecteurs  $r$  et  $s$  ne peuvent être le résultat d'un simple calcul matriciel linéaire.

Si l'on ajuste la matrice initiale  $X_0$  successivement sur les lignes et sur les colonnes, les produits des multiplicateurs de lignes et de colonnes convergent vers les éléments des vecteurs  $r$  et  $s$ .

La fonction de distance que l'on minimise a la forme suivante:

$$d(X^*, X_0) = \sum_i \sum_j x_{ij}^* \cdot \ln \left\{ \frac{x_{ij}^*}{(e \cdot x_{ij}^0)} \right\}$$

Il ne s'agit clairement pas d'une simple fonction de distance euclidienne<sup>5</sup>.

La relation fonctionnelle de la méthode RAS maintient la stricte positivité et les valeurs nulles des éléments correspondants de  $X^*$  et de  $X_0$ :

$$r_i \cdot 0 \cdot s_j = 0$$

Cette dernière caractéristique est d'une grande importance lors de l'utilisation de la méthode RAS pour l'équilibrage des tableaux entrées-sorties.

---

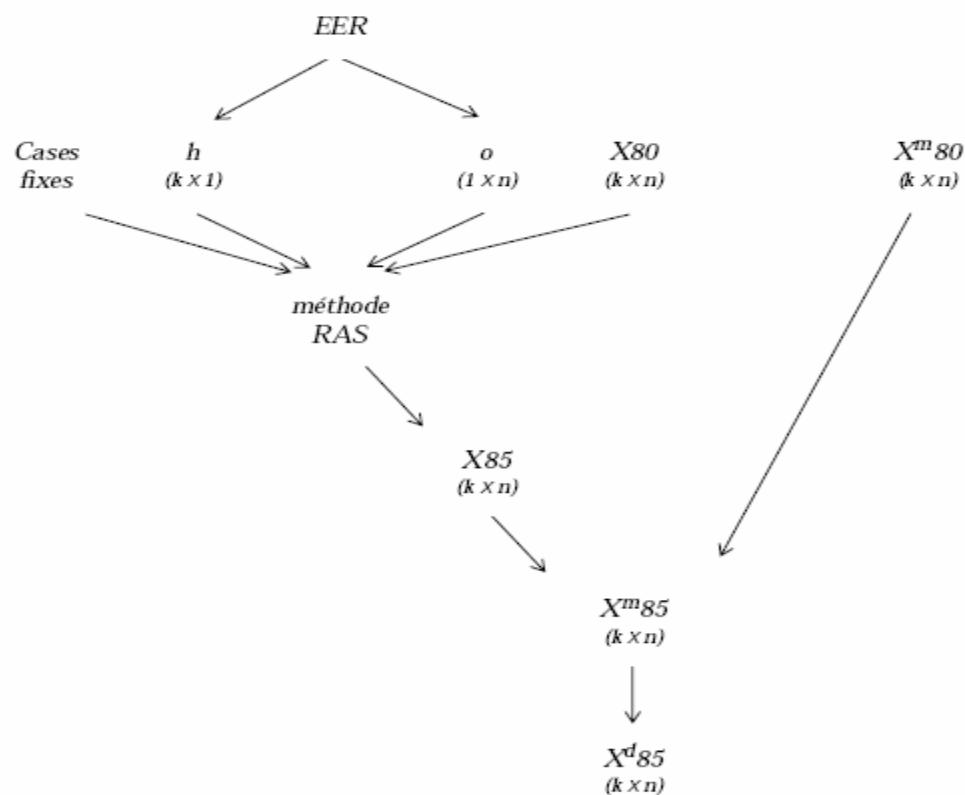
<sup>5</sup> Bacharach M. (1969), "Biproportional Matrices and Input-Output Change", Cambridge University Press, pp.42-58 et pp.75-86.

## Démarche de calcul du TEI<sup>6</sup>

### Graphique 3 : Illustration du Calcul du TEI

#### Calcul du TEI 1985

---



#### Sigles

h: vecteur représentant le total des sorties intermédiaires

o: vecteur représentant le total des entrées intermédiaires

k: nombre de produits

n: nombre de branches

EER: tableau des équilibres emplois-ressources

X: matrice des échanges intermédiaires ou d'absorption

Xm: matrice d'absorption d'origine importée

Xd: matrice d'absorption d'origine nationale

Les valeurs qui figurent entre parenthèses sous les différentes matrices (ou éventuellement vecteurs) donnent la dimension des matrices (ou vecteurs) en question.

---

<sup>6</sup> Démarche inspirée du document : Tableau Entrées-Sorties 1985 : une analyse des structures économiques de la Belgique ; L.Avonds, J.floridor, A.Gilot, C.Hambye, D.Rase, K,Verstegen, Octobre 1998.

## **II-DEUXIEME PARTIE**

## II-1 METHODOLOGIE DE CALCUL DES COEFFICIENTS TECHNIQUES ROBUSTES

Cette partie rappelle la méthodologie élaborée dans le cas du Sénégal pour la détermination des coefficients techniques robustes, ainsi que la méthode d'évaluation des performances du modèle.

### II-1-1 Le schéma synthétique de la ventilation

L'étude devrait se faire sur les séries disponibles de coefficients techniques. Le TES du Sénégal étant composé de 41 branches et 41 produits, on assiste tout naturellement à une augmentation drastique du nombre de séries à étudier (ici 1681). Pour pallier ce travail fastidieux, le schéma de ventilation ci-dessous a été élaboré dans le but d'expliquer intuitivement les différents mécanismes qui concourent au calcul de ces coefficients techniques.

Dans cette section, la structure des coefficients techniques calculés ci-dessus ( $\gamma_{\bullet,j}$ )<sup>7</sup> est ventilée avec la structure de la part des produits dans chaque branche ( $\beta_{i\bullet}$ ).

#### Définition des termes

$T_j$  représente le total des consommations intermédiaires par branches,  $T_j = \sum_{i=1}^{41} CI_{ij}$

$T_{i\bullet} = \text{Médiane}\{\text{distribution}(CI_{ij})\}$  et  $TT_{i\bullet} = \sum_{j=1}^{41} \text{Médiane}\{\text{distribution}(CI_{ij})\}$

$\beta_{i\bullet} = \frac{T_{i\bullet}}{TT_{i\bullet}} = \frac{\text{Médiane}\{\text{distribution}(CI_{ij})\}}{\sum_{j=1}^{41} T_{i\bullet}}$  représente la structure de la part de chaque produit  $i$  dans la branche  $j$ .

$\gamma_{\bullet,j} = \frac{T_j}{PDB_j} = \frac{\sum_{i=1}^{41} CI_{ij}}{PDB_j}$  structure des coefficients techniques, et PDB la production disponible de la branche  $j$

$$\alpha_{ij} = \beta_{i\bullet} * \gamma_{\bullet,j}$$

$\alpha_{ij}$  représente la structure estimée du coefficient technique du produit  $i$  dans la branche  $j$

<sup>7</sup> Pour le calcul des médianes et de la moyenne, l'indice de temps a été omis pour des raisons de lisibilité du document.

### II-1-2 Détermination de la structure des $\beta_{i\cdot}$

La détermination des  $\beta_{i\cdot}$  se fait à l'aide d'un TES médian calculée sur la distribution des 25 années de consommations intermédiaires stockées dans chaque case du TES.

La détermination de ce TES médian implique le calcul des quantités

$$T_{i\cdot} = \text{Médiane} \left\{ \text{distribution} (CI_{ij}) \right\}$$

Ensuite, la somme du produit  $i$  utilisée par chaque branche est calculée ( $TT_{i\cdot}$ )

$$TT_{i\cdot} = \sum_{j=1}^{41} \text{Médiane} \left\{ \text{distribution} (CI_{ij}) \right\}$$

Enfin la structure des  $\beta_{i\cdot}$  représente la part médiane de chaque produit  $i$  utilisée dans toutes les branches, de telle sorte que:

$$\beta_{i\cdot} = \frac{T_{i\cdot}}{TT_{i\cdot}} = \frac{\text{Médiane} \left\{ \text{distribution} (CI_{ij}) \right\}}{\sum_{j=1}^{41} T_{i\cdot}}, \text{ et par construction } \sum_{j=1}^{41} \beta_{ij} = 1 .$$

### II-1-3 Détermination de la structure des $\gamma_{\cdot j}$

La détermination de la structure des coefficients techniques se fait en utilisant la formule suivante :

$$\gamma_{\cdot j} = \frac{T_{\cdot j}}{PDB_j} = \frac{\sum_{i=1}^{41} CI_{ij}}{PDB_j} \rightarrow \sum_j \gamma_{\cdot j} \neq 1$$

$\gamma_{\cdot j}$  est le vecteur qui stocke les coefficients techniques par branches.

Le schéma explicatif du processus est illustré ci-dessous :

**Graphique 4 : Schéma de la ventilation**

	Branche 1	...	Branche j	...	Branche 41	TOTAL	
Produit 1							
...							
...							
...							
...							
Produit i			$\alpha_{ij}$			$T_{i\bullet}$	$\beta_{i\bullet}$
...							
...							
...							
...							
Produit 41							
TOTAL CI			$T_{\bullet j}$				
			$\gamma_{\bullet j} = \frac{T_{\bullet j}}{PDB_j} = \frac{\sum_{i=1}^{41} CI_{ij}}{PDB_j}$				

### II-1-4 Le processus d'élaboration des coefficients techniques robustes

Il est décliné en trois étapes.

La première étape consiste à étudier chaque branche qui requiert l'analyse de la série des coefficients techniques sous le prisme des techniques économiques standards disponibles au niveau de la littérature empirique. Chaque série est décomposée sous la forme (pour des raisons de simplicité):

$$Y_t = Tendance + Cycle + \varepsilon_t \quad \varepsilon_t \rightarrow N(0, \sigma^2) \quad \varepsilon_t \text{ étant un bruit blanc}$$

A cet effet, une batterie de tests et de filtres statistiques notamment le filtre Hodrick Prescott (HP)<sup>9</sup>, la méthode des Holt Winters (HW)<sup>10</sup>, sont utilisées. Pour chaque série, le meilleur filtre est déterminé en utilisant les tests RMSE (Root Mean Square Error : erreur quadratique moyenne) et le coefficient de THEIL. Ensuite, la moyenne et la médiane sont calculées pour chaque série. Enfin, l'intervalle de confiance de la médiane ou de la moyenne qui est  $Me \pm 2RMSE$ ,  $Moy \pm 2RMSE$  est déterminé, aussi bien pour les consommations intermédiaires que pour les productions disponibles.

<sup>8</sup> Cette décomposition inclut les processus TS et DS

<sup>9</sup> Voir Note technique en annexe

<sup>10</sup> Voir annexe



Dans la seconde étape, le processus ci-dessus élaboré pour les CI est itéré.

Enfin, la troisième étape consiste à élaborer les coefficients techniques robustes avec la formule suivante :

$$a- \alpha_{ij} = \frac{Me(\hat{CI}_{ij})}{Me(P\hat{D}Br_j)} \pm \frac{RMSE(\hat{CI}_{ij})}{RMSE(P\hat{D}Br_j)} \text{ (voir encadré 3 en annexe)}$$

$$b- \beta_{ij} = \frac{Moyenne (\hat{CI}_{ij})}{Moyenne (P\hat{D}Br_j)} \pm \frac{RMSE (CI_{ij})}{RMSE (P\hat{D}Br_j)}$$

Il y a j branches  $j \in [10 - 410]$ <sup>11</sup>. Les valeurs utilisées sont la production disponible et la consommation intermédiaire totale de la branche. Soit  $PDBr_j$ , la production disponible de la branche j,  $CIBr_j$ , la consommation intermédiaire totale de la branche j :

On définit le coefficient technique  $\alpha_{ij}$  du produit i dans la branche j comme :

$$\alpha_{ij} = \frac{CI_{ij}}{PDBr_j}$$

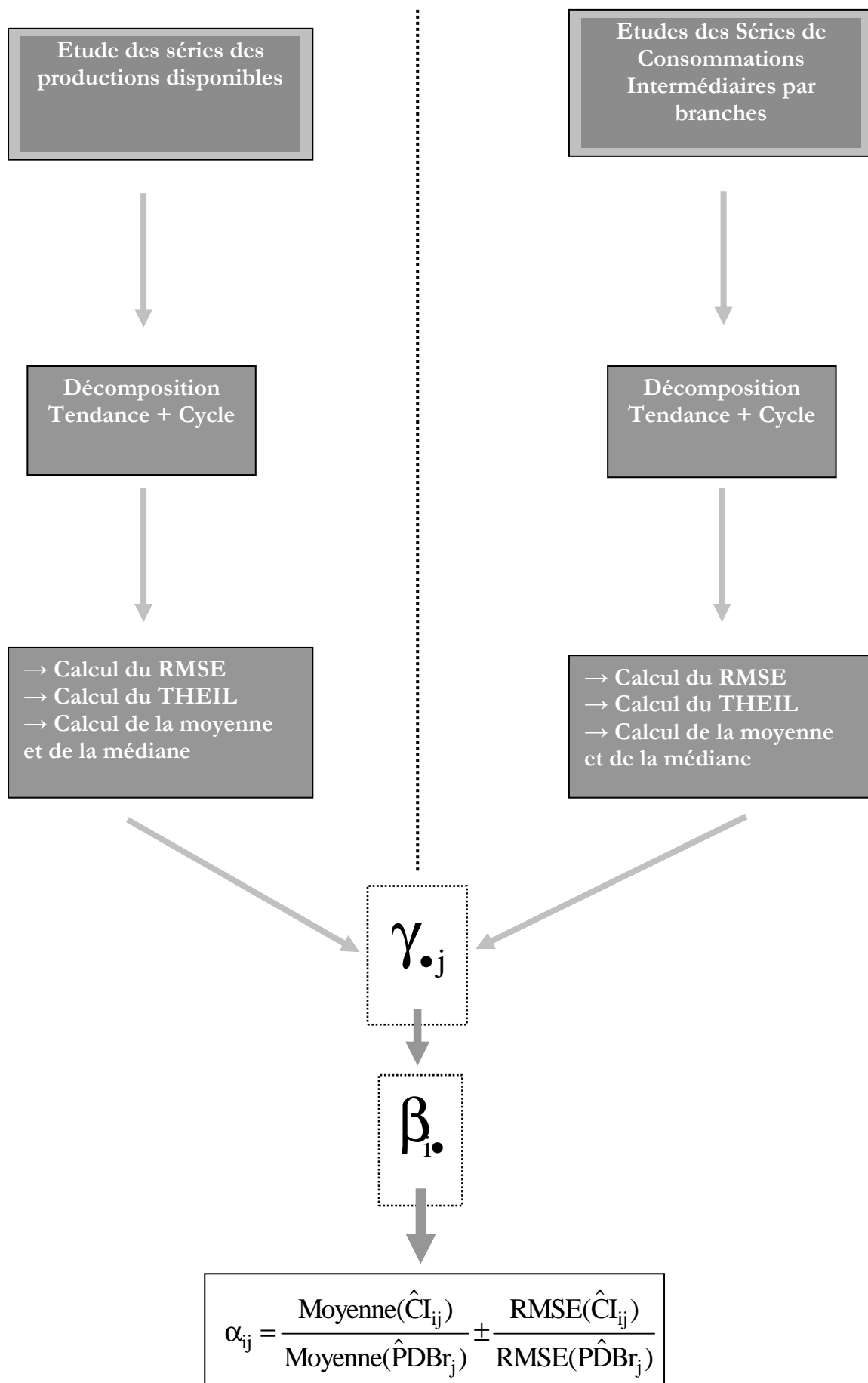
$CI_{ij}$  représente la consommation intermédiaire en produit i de la branche j.

$\hat{CI}_{ij}$  représente la consommation intermédiaire estimée en produit i de la branche j.

---

<sup>11</sup> Les branches 10 et 410 proviennent de la nomenclature NAEMAS

Schématiquement, le processus d'élaboration des coefficients techniques est indiqué ci-dessous :



## II-2 LA DETERMINATION DE L'INTERVALLE DE CONFIANCE

Cette section traite de la détermination des intervalles. Le premier intervalle est lié à une fonction pivotale de la moyenne, tandis que la seconde est fondée sur des hypothèses empiriques.

### II-2-1 Construction d'un intervalle de confiance à partir d'un pivot

On considère un échantillon  $\{Y_1, \dots, Y_n\} \rightarrow \text{iid} \sim N(m_y, \sigma_y^2)$  et  $\{X_1, \dots, X_n\} \rightarrow \text{iid} \sim N(m_x, \sigma_x^2)$

On démontre (voir encadré 3) que  $Z = \frac{X}{Y}$  suit sous certaines conditions une loi normale.

$\{Z_1, \dots, Z_n\} \rightarrow \text{iid} \sim N(m, \sigma^2)$

Loi de l'estimateur de m.

On démontre que la moyenne Z est estimée séparément par le ratio  $\hat{Z} = \frac{\hat{X}}{\hat{Y}}$

Avec  $\hat{X} = \frac{1}{n} \sum_{i=1}^n X_i$  représentant la moyenne de la distribution des  $X_i$

Et  $\hat{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$  représentant la moyenne de la distribution des  $Y_i$

La variance  $\sigma^2$  connue est calculée comme étant égale à  $\sigma^2 = m^2 * \left[ \frac{\sigma_x^2}{m_x^2} + \frac{\sigma_y^2}{m_y^2} \right]$

On définit alors la statistique pivotale

$$\frac{Z - \mu}{(m) * \sqrt{\frac{\sigma_y^2}{m_y^2} + \frac{\sigma_x^2}{m_x^2}}} \rightarrow N(0, 1)$$

### Construction de l'intervalle de confiance

On note  $u_p$  le quantile d'ordre p de la loi normale.

La loi de Z implique que

$$\text{Pr} \left\{ u_{\alpha/2} \leq \frac{Z - \mu}{m * \sqrt{\frac{\sigma_y^2}{m_y^2} + \frac{\sigma_x^2}{m_x^2}}} \leq u_{1-\alpha/2} \right\}$$

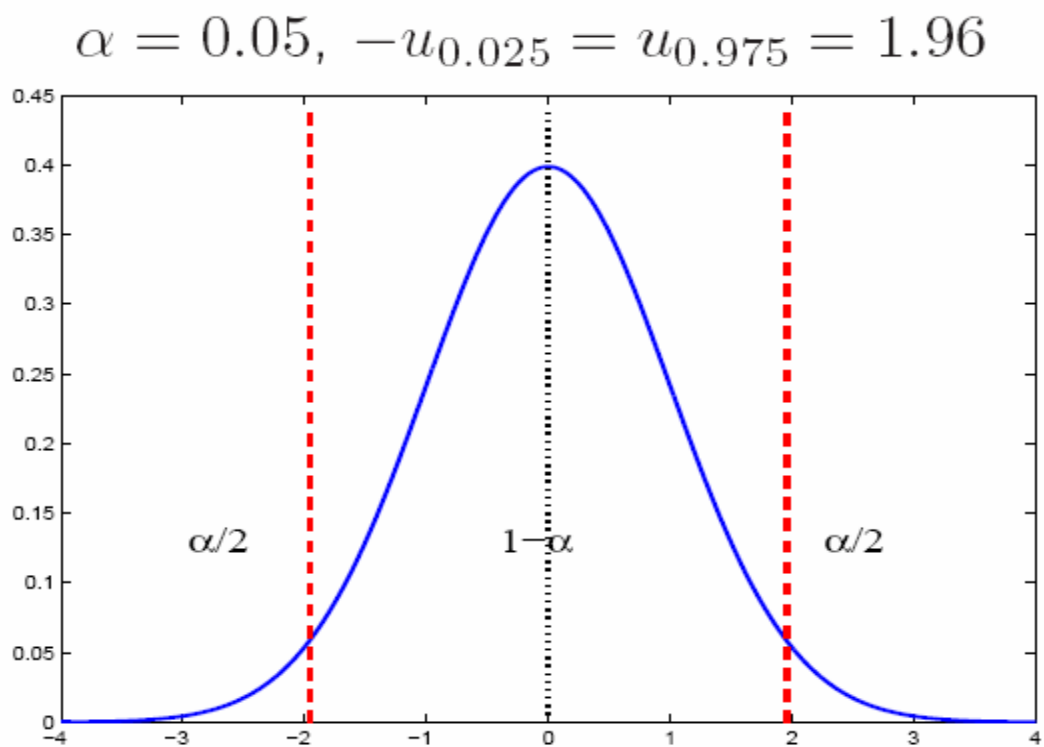
$$\text{soit } \Pr \left\{ \overbrace{\hat{Z} + u_{\alpha/2} * m * \sqrt{\frac{\sigma_y^2}{m_y^2} + \frac{\sigma_x^2}{m_x^2}}}^{B_1} \leq \mu \leq \overbrace{\hat{Z} + u_{1-\alpha/2} * m * \sqrt{\frac{\sigma_y^2}{m_y^2} + \frac{\sigma_x^2}{m_x^2}}}^{B_2} \right\} = 1 - \alpha$$

L'intervalle de confiance est alors

$$IC_{1-\alpha}(\alpha) = \left[ \hat{Z} \pm u_{1-\alpha/2} * m * \sqrt{\frac{\sigma_x^2}{m_x^2} + \frac{\sigma_y^2}{m_y^2}} \right]$$

$$\text{et la largeur de l'intervalle est } B_2 - B_1 = \left[ 2 * u_{1-\alpha/2} * m * \sqrt{\frac{\sigma_x^2}{m_x^2} + \frac{\sigma_y^2}{m_y^2}} \right]$$

**Graphique 5 : Courbe d'une loi normale et illustration des bornes de l'intervalle de confiance**



## II-2-2 Construction d'un intervalle de confiance empirique

Deux intervalles de confiance ont été déterminés. L'un centré sur la moyenne et l'autre sur la médiane. De façon explicite, le premier intervalle est de la forme<sup>12</sup> :

$$\left[ \frac{\text{Moyenne}(\hat{CI})}{\text{Moyenne}(\hat{PDBBranche})} - \frac{\text{RMSE}(\hat{CI})}{\text{RMSE}(\hat{PDBBranche})}; \frac{\text{Moyenne}(\hat{CI})}{\text{Moyenne}(\hat{PDBBranche})} + \frac{\text{RMSE}(\hat{CI})}{\text{RMSE}(\hat{PDBBranche})} \right]$$

Le second intervalle de confiance est de la forme suivante :

$$\left[ \frac{\text{Médiane}(\hat{CI})}{\text{Médiane}(\hat{PDBBranche})} - \frac{\text{RMSE}(\hat{CI})}{\text{RMSE}(\hat{PDBBranche})}; \frac{\text{Médiane}(\hat{CI})}{\text{Médiane}(\hat{PDBBranche})} + \frac{\text{RMSE}(\hat{CI})}{\text{RMSE}(\hat{PDBBranche})} \right]$$

Les résultats de l'analyse montrent que les estimateurs, pris individuellement, ne sont pas assez robustes. Les écarts entre les valeurs estimées et celles prédites sont assez importants. Un tel résultat montre que les intervalles de confiance sont assez larges et les coefficients techniques estimés peu fiables. Pour remédier à ce problème, les changements suivants ont été apportés.

Au lieu de considérer un intervalle de confiance centré uniquement sur la moyenne ou la médiane, l'intervalle suivant a été élaboré :

Borne inférieure=

$$\text{Max} \left[ \left( \frac{\text{Moyenne}(\hat{CI})}{\text{Moyenne}(\hat{PDBBranche})} - \frac{\text{RMSE}(\hat{CI})}{\text{RMSE}(\hat{PDBBranche})}; \frac{\text{Médiane}(\hat{CI})}{\text{Médiane}(\hat{PDBBranche})} - \frac{\text{RMSE}(\hat{CI})}{\text{RMSE}(\hat{PDBBranche})} \right) \right]$$

De même pour la borne supérieure, nous considérons l'estimateur suivant :

$$\text{Borne supérieure} = \text{Min} \left[ \left( \frac{\text{Moyenne}(\hat{CI})}{\text{Moyenne}(\hat{PDBBranche})} + \frac{\text{RMSE}(\hat{CI})}{\text{RMSE}(\hat{PDBBranche})}; \frac{\text{Médiane}(\hat{CI})}{\text{Médiane}(\hat{PDBBranche})} + \frac{\text{RMSE}(\hat{CI})}{\text{RMSE}(\hat{PDBBranche})} \right) \right]$$

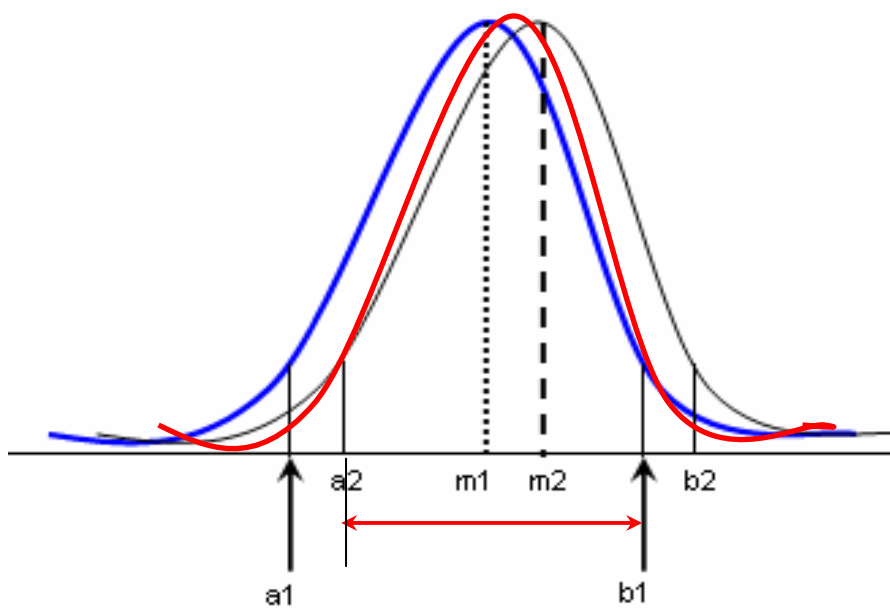
Cette méthode a surtout pour avantage de réduire la largeur de l'intervalle de confiance.

Les résultats des investigations sont illustrés dans le tableau 6.

Le graphique ci-dessous illustre la méthode utilisée.

<sup>12</sup> Pour plus de détails dans la démonstration, voir les manuels de mathématique sur les mélanges de lois

Graphique 6 : Illustration des lois suivies par les deux estimateurs



## **TROISIEME PARTIE**

### III-1 LES PERFORMANCES DU MODELE

Les performances globales du modèle sont évaluées à partir de deux critères. Le premier critère consiste à construire des indicateurs de mesure de l'écart entre les valeurs simulées par le modèle et les valeurs observées de quelques variables endogènes clefs. L'indicateur retenu est le coefficient d'inégalité de Theil. Il peut être interprété comme l'erreur relative en moyenne par année. Plus la valeur du coefficient de Theil est faible, plus l'erreur de prévision du modèle est faible. Le coefficient de Theil peut être décomposé en trois éléments dont la somme est égale à l'unité :

- la part de l'erreur de prévision systématique imputable à la différence de la moyenne des valeurs simulées et observées ;
- la part de l'erreur de prévision due à la structure du modèle ;
- la part de l'erreur de prévision qui est de nature résiduelle.

Le deuxième critère d'appréciation des performances globales du modèle est l'erreur quadratique moyenne (RMSE : Root Mean Square Error). Il est un indicateur définissant le niveau de précision du modèle.

Root Mean Squared  
Error

$$\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}$$

Theil Inequality  
Coefficient

$$\frac{\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} \hat{y}_t^2 / h} + \sqrt{\sum_{t=T+1}^{T+h} y_t^2 / h}}$$

Avec  $y_t$  : la valeur de la variable

$\hat{y}_t$  : la valeur estimée de la variable.



## III-2 ESTIMATION

Cette partie passe en revue l'estimation du TES robuste et le commentaire des résultats obtenus.

### III-2-1 Estimation du TES robuste

Dans les sections précédentes, l'estimation des coefficients techniques robustes a été faite avec une technique appropriée. Cependant, il faut noter que ces coefficients techniques représentent en réalité les taux par branches des consommations intermédiaires. Le dernier point de cette étude consiste à déterminer une structure de poids qui permet de ventiler les coefficients techniques déjà calculés. De ce fait, la structure stabilisée est celle des coefficients techniques des branches. Une telle ventilation a été faite en utilisant comme estimateur la médiane.

Dans un premier temps, la part de chaque produit dans la production disponible de la branche est calculée pour chaque année sur la série des 25 années de TES. Ensuite, sur les parts ainsi calculées, la médiane est déterminée, ce qui permet l'obtention de la structure qui ventile les coefficients techniques. Les résultats des calculs sont illustrés dans le tableau 4.

### III-2-2 Commentaire des résultats du tableau 7 ( voir annexe)

Dans le secteur primaire, les calculs de l'erreur montrent une très bonne adéquation entre les estimations et la structure réelle du TES de 2001 à 2004. Le minimum de l'erreur calculé (**2,77%**), est constaté au niveau de la pêche ; alors que le maximum de l'erreur calculé (**9%**) provient de l'agriculture vivrière. En moyenne l'erreur dans le secteur primaire hors « extractives » est de **5,06%**.

L'explication du niveau d'erreur élevé au niveau de la pêche, vient principalement du fait que les données ne sont pas bien captées. Il faut aussi prendre en compte le fait que la structure de la filière de la pêche n'est pas totalement cernée. En effet, la structure des coefficients utilisés dans les comptes doit être mise à jour à moyen terme pour tenir compte des nouvelles réalités de la filière.

Les estimations dans la structure de TES, ont montré qu'avec une plage d'erreur comprise entre **7,30%** (pour la branche « Fabrication de cuir ») et **16,74%** (pour la branche « travail de grains, fabrication de produits »), le secteur secondaire a le niveau d'erreur le plus élevé observé (avec une moyenne de **12,49%**). Cependant, les résultats obtenus montrent une bonne adéquation entre la structure des TES et les estimations faites.

Le niveau d'erreur élevé en moyenne dans le secteur moderne interpelle naturellement sur la qualité des informations reçues d'une part et ou sur les méthodologies d'estimation utilisées d'autre part. Il est paradoxal de constater que le secondaire enregistre le niveau d'erreur le plus élevé dans la mesure ou dans ce secteur l'information devrait être plus fine et largement disponible.

L'explication de l'erreur dans le secteur secondaire pourrait provenir de l'hétérogénéité fortement marquée dans les structures de production des différentes branches, ainsi que du poids démographique du secondaire

qui représente à peu près 59% du total des branches. D'autre part, il est possible que la structure des coefficients techniques soit plus sensible aux impacts technologiques, du fait des investissements en matériel effectués par les entreprises. Ainsi, la structure pourrait changer à court terme, alors que le modèle prévoit une stabilité sur cet horizon.

L'erreur calculé sur les estimations dans la structure du TES dans le secteur tertiaire affiche une moyenne de **8,11%**, avec un minimum de **0,01%** au niveau de la branche « activités à caractère collectif ou personnel » et un maximum de **16,85%** au niveau de la branche « services d'hébergement et restauration ».

Les estimations faites corroborent la structure du TES de ces branches sur la période 2001 à 2004.

Au niveau du tertiaire, une attention particulière doit être accordée à la branche « services d'hébergement et restauration », puisque le niveau d'erreur observé est assez important.

Sur un autre registre, une simulation est effectuée pour vérifier si les intervalles de confiance élaborées pour chaque coefficient sont réalistes. Cette simulation est effectuée en calculant la structure réelle des TES de 2001 à 2004 d'une part, et en vérifiant d'autre part si chaque coefficient technique « réel » calculé se trouve dans son intervalle de confiance.

L'observation du tableau 6, permet de répondre par l'affirmative. En effet, la tableau 6 montre que tous les coefficients techniques sont à l'intérieur des intervalles de confiances estimées. Pour, plus de finesse dans les résultats, nous avons utilisé la structure des coefficients techniques du TES semi définitif de 2005. La structure de ce TES de 2005, révèle que les intervalles de confiances sont stables.

### III-3 SIMULATIONS ET COMPARAISONS

#### III-3-1 Choix de l'estimateur

Pour déterminer l'estimateur le plus adéquat, un test simple de résidus est effectué en confrontant les estimations faites aux données réelles. Le tableau ci-dessous illustre les résultats et démontrent sans ambiguïté que la moyenne est plus adéquate que la médiane.

**Tableau 1:** Test d'adéquation et de performance des estimations sur les deux estimateurs retenus

	Coefficient de Theil		Erreur Quadratique moyenne	
	Médiane	Moyenne	Médiane	Moyenne
Structure du TES 2001	7,37%	2,90%	15,11%	5,73%
Structure du TES 2002	7,34%	2,90%	15,06%	5,75%
Structure du TES 2003	7,32%	3,05%	15,07%	6,05%
Structure du TES 2004	7,32%	3,05%	15,07%	6,05%

L'analyse du tableau ci-dessus montre que le coefficient de Theil est plus faible lorsque les estimations sont faites avec la moyenne soit environ 3%, alors que dans le cas de la médiane le coefficient est de l'ordre de 7,3% en moyenne.

De même, la simulation effectuée avec l'erreur quadratique moyenne montre que l'erreur obtenue en utilisant la médiane est trois fois plus élevée que celle faite en utilisant la moyenne.

Il ressort de ces tests que la moyenne est l'estimateur le plus adéquat dans l'estimation des coefficients techniques robustes.

### III-3-2 Simulation d'une augmentation de 10% du prix du baril du pétrole.

Tableau 2 : Simulation d'une augmentation du prix de l'énergie

Coefficients techniques	MODELE	2004 Volume	2005 Valeur	2005 Volume	Ecart
<b>010 agriculture vivriere</b>	<b>2,43%</b>	<b>2,21%</b>	<b>2,78%</b>	<b>2,57%</b>	<b>0,21%</b>
020 agriculture industrielle ou d'exportation	1,59%	1,44%	1,81%	1,68%	0,14%
030 élevage et chasse	1,65%	1,50%	1,89%	1,75%	0,14%
040 sylviculture, exploitation forestiere,	0,22%	0,20%	0,26%	0,24%	0,02%
050 peche	1,00%	0,91%	1,14%	1,06%	0,09%
060 activites extractives	0,58%	0,53%	0,66%	0,61%	0,05%
<b>070 Transformation et conservation de viande, poisson</b>	<b>2,33%</b>	<b>2,12%</b>	<b>2,66%</b>	<b>2,46%</b>	<b>0,20%</b>
080 fabrication de corps gras alimentaires	1,52%	1,38%	1,74%	1,61%	0,13%
090 travail de grains, fabrication de produits	1,15%	1,05%	1,32%	1,22%	0,10%
100 fabrication de produits alimententaires a	0,96%	0,87%	1,10%	1,02%	0,08%
110 fabrication de sucre, transformation	0,49%	0,44%	0,56%	0,52%	0,04%
120 fabrication de produits alimentaires n.c.a	0,57%	0,52%	0,66%	0,61%	0,05%
130 fabrication de boissons	0,21%	0,19%	0,23%	0,22%	0,02%
140 fabrication de produits a base de tabac	0,23%	0,21%	0,27%	0,25%	0,02%
150 egrenage de coton et fabrication des textiles	1,03%	0,94%	1,18%	1,09%	0,09%
160 fabrication du cuir; fabrication	0,06%	0,05%	0,07%	0,06%	0,01%
170 travail du bois et fabrication d'articles	0,32%	0,29%	0,37%	0,34%	0,03%
180 fabrication de papier, carton,	0,41%	0,38%	0,47%	0,44%	0,04%
190 raffinage petrole, cokefaction,	1,00%	0,91%	1,14%	1,05%	0,09%
200 fabrication de produits chimiques	0,96%	0,87%	1,10%	1,01%	0,08%
210 fabrication de produits en caoutchouc	0,26%	0,23%	0,29%	0,27%	0,02%
220 fabrication de verre, poterie	0,33%	0,30%	0,38%	0,35%	0,03%
230 metallurgie, fonderie, fabrication	0,36%	0,33%	0,42%	0,38%	0,03%
240 fabrication de machines	0,07%	0,07%	0,08%	0,08%	0,01%
260 construction de materiels de transports	0,04%	0,03%	0,04%	0,04%	0,00%
270 fabrication de produits divers	0,46%	0,41%	0,52%	0,48%	0,04%
280 electricite, gaz et eau	1,30%	1,18%	1,48%	1,37%	0,11%
<b>290 construction</b>	<b>3,03%</b>	<b>2,75%</b>	<b>3,46%</b>	<b>3,20%</b>	<b>0,26%</b>
<b>300 commerce</b>	<b>6,34%</b>	<b>5,77%</b>	<b>7,25%</b>	<b>6,70%</b>	<b>0,55%</b>
310 services de la reparation	0,75%	0,68%	0,86%	0,79%	0,07%
320 services d'hebergement et restauration	1,54%	1,40%	1,76%	1,63%	0,13%
<b>330 transports</b>	<b>2,86%</b>	<b>2,60%</b>	<b>3,27%</b>	<b>3,02%</b>	<b>0,25%</b>
340 postes et telecommunications	1,09%	0,99%	1,25%	1,15%	0,09%
350 services financiers	0,97%	0,88%	1,11%	1,02%	0,08%
360 activites immobilieres	1,56%	1,42%	1,78%	1,64%	0,14%
370 activites des services aux entreprises	1,29%	1,17%	1,47%	1,36%	0,11%
<b>380 activites d'administration publique</b>	<b>2,81%</b>	<b>2,56%</b>	<b>3,21%</b>	<b>2,97%</b>	<b>0,24%</b>
390 education et formation	1,71%	1,55%	1,95%	1,80%	0,15%
400 activites de sante et action sociale	0,58%	0,52%	0,66%	0,61%	0,05%
410 activites a caractere collectif ou personnel	0,81%	0,74%	0,93%	0,86%	0,07%
	<b>46,88%</b>	<b>42,63%</b>	<b>53,59%</b>	<b>49,51%</b>	<b>4,08%</b>

Dans cette section, on simule une augmentation du prix du baril du pétrole sur la structure des coefficients techniques du TES robuste estimé.

Entre 2004 et 2005, le prix du baril est passé environ de 65\$ à 70\$ US, soit une augmentation de 7,7%. Cette augmentation de 8% environ au niveau des prix se répercute au niveau des coefficients techniques de l'énergie avec un impact global de 16,14%.

L'observation du tableau 2 met en évidence les principales branches qui ont le plus subi la flambée des prix du pétrole durant cette période. La flambée des prix a été le plus défavorable en général aux branches qui sont les plus utilisatrices de ressources énergétiques. Au nombre de ces branches, il faut noter :

- ✓ La branche « Agriculture vivrière », au niveau du secteur primaire connaît une augmentation de ces prix de 0,21%. Par rapport à 2004, les coefficients techniques ont subi une poussée de 0,36%.
- ✓ Au niveau du secondaire, deux branches se distinguent. La branche « Transformations et Conservations de viandes et poissons » ainsi que la branche « Construction ». Les prix au niveau de chacune de ces branches sont respectivement de 0,20% et 0,26%. Par rapport à 2004, on note respectivement, une augmentation des coefficients techniques de 0,34% et 0,45%.
- ✓ Dans le secteur tertiaire, trois branches ont le plus subi le choc des prix : il s'agit du « commerce » avec une augmentation des prix de 0,55%, et une évolution des coefficients techniques de 0,93% par rapport à 2004. Ensuite vient « les Transports », avec une augmentation des prix de 0,25%, avec une évolution des coefficients techniques de 0,42%. Enfin, les « Activités d'administration publiques », où l'on note une poussée des prix de 0,24%, avec une évolution des coefficients techniques de 0,41% par rapport à 2004.

Les résultats de cette simulation renseignent sur les capacités des structures estimées à reproduire les faits réels de l'économie. Ainsi, les branches qui ont le plus subi le choc de la flambée des prix sont mises en évidence, de même que l'ampleur de l'impact sur ces branches. Il apparaît donc clairement, bien qu'il puisse exister des pistes d'amélioration dans le futur que l'estimation de la structure des coefficients techniques estimées, et corollairement les intervalles de confiance sont assez stables.

## LIMITES DE L'ETUDE ET CONCLUSION

La construction du TES et ses applications sont le fruit d'un grand nombre d'investigations statistiques souvent fastidieuses et de développements méthodologiques parfois très complexes. Afin de rester lisible, cette étude s'est limitée à décrire les grandes lignes des estimations réalisées, de même qu'elle limite au minimum indispensable le volet théorique, laissant au lecteur intéressé le soin de se référer aux notes techniques sur lesquelles se basent les estimations, ou les références bibliographiques sur la technique des modèles entrées-sorties.

Le tableau entrées-sorties peut, fournir un grand nombre d'informations sur la structure des coûts, des débouchés et l'interdépendance entre les branches. Il peut aussi sur un autre registre, être utilisé comme instrument pour mesurer les effets directs et indirects de modifications de la demande finale ou de la production d'une branche sur le reste de l'économie.

Des modifications de production au niveau d'une branche ont bien entendu des répercussions sur les autres branches, que ce soit "en amont", c-à-d sur les branches fournisseurs, ou "en aval", c-à-d sur les branches utilisatrices.

La robustesse des coefficients techniques, ne signifie guère le fait qu'ils soient statistiques. La robustesse appréhendée ici, est le fait que les coefficients techniques soient chacun compris dans un intervalle précis. Ainsi, au sens de notre étude, la robustesse des coefficients techniques traduit le fait que malgré les fluctuations et les chocs subis par l'économie pendant 25 ans, les coefficients ne peuvent fluctuer indépendamment.

Les résultats obtenus mettent en évidence que les coefficients techniques sont tous inclus dans les intervalles de confiance déterminés. Ce qui corroborent les hypothèses de base de nos modèles.

Toutefois malgré la confirmation des hypothèses de base sur la robustesse des coefficients techniques calculés, il est indéniable que l'étude comporte des limites.

Les limites de cette étude sont de trois types.

- 1- dans un premier temps la dimension temporelle de l'échantillon considéré est assez courte (25 années). Ceci peut avoir pour conséquence, la non robustesse des estimations économétriques effectuées. En effet il est établi que la robustesse des tests statistiques est prise en distance finie.
- 2- L'étude se concentre sur l'estimateur d'un ratio. Les lois suivis par le numérateur et le dénominateur étant des lois normales, il est connu que la loi suivi par le ratio est une loi de Cauchy (dans le cas ou le numérateur et le dénominateur suivent des lois centrées réduites), ou une loi plus complexe encore le cas échéant. En conséquence, les approximations faites pour approcher la distribution du ratio par une

loi normale, donnent des intervalles de confiance très large. Une des recommandations devrait être, de travailler directement sur une série de coefficients techniques déjà calculés.

- 3- Enfin, l'intervalle de confiance élaboré est basé sur la moyenne et la médiane. Alors que la fonction pivotale de la loi suivie par cet estimateur (qui a simultanément les propriétés de la médiane et de la moyenne) est quasi inconnue.

Les recommandations générales portent sur les points suivants :

- 1- la mise à jour des coefficients techniques une fois les séries de comptes trimestriels disponibles. Cette mise à jour est indispensable puisqu'elle permettra de stabiliser les coefficients techniques.
- 2- L'utilisation effective de la série des coefficients techniques élaborées par le bureau de la comptabilité nationale de l'ANSD et des institutions utilisant les statistiques à l'instar de la DPEE, de la Direction du Commerce Extérieur, la CSLCP, etc.
- 3- Une publication de l'étude dans les colonnes officielles de l'agence, pour une diffusion large des résultats obtenus.

## BIBLIOGRAPHIE

- Akhabbar Amanar [2006], « Machine de paix, machine de guerre : l'analyse *input-output* comme modèle pour l'intervention publique », Communication au Colloque Etat Social du MATISSE (Université Paris 1), Paris, 11-13 septembre.
- Barengolts M. [1928], "Capacity of the industrial market in the USSR", in Spulber [1964b], traduit du russe, "Emkost promyshlennogo rynka v SSSR", *Planovoe khozjaistvo*, juillet 1928, n°7.
- Belykh A.A. [1989], "A note on the origins of input-output analysis and the contribution of the early soviet economists: Chayanov, Bogdanov and Kristman", *Soviet Studies*, 41(3), p. 426-429.
- Belykh A.A. [1989], "A note on the origins of input-output analysis and the contribution of the early soviet economists: Chayanov, Bogdanov and Kristman", *Soviet Studies*, 41(3), p. 426-429.
- Boukharine Nicolas et Preobrajensky Eugène [1919], *ABC du communisme*, Paris, Maspero (1968).
- Boukharine Nicolas I. [1928], "Notes of an economist at the beginning of a new economic year", in Spulber [1964b], traduit du russe.
- Clark D.L [1984], "Planning and the real origins of input-output analysis", *Journal of Contemporary Asia*, 14(4), p.408-415.
- Clark D.L [1984], "Planning and the real origins of input-output analysis", *Journal of Contemporary Asia*, 14(4), p.408-415.
- Domar (Domashevitski) Evsey David [1957], "A soviet model of growth", repris dans E.D. Domar, *Essays in the theory of economic growth*, New-York: Oxford University Press.
- Granik D. [1955], *Management of the industrial firm in the USSR*, New-York, Columbia University Press.
- Jasny Naum [1962a], "The Russian economic 'balance' and input-output analysis: a historical comment", *Soviet Studies*, 14(1), p.75-80.
- Jasny Naum [1962a], "The Russian economic 'balance' and input-output analysis: a historical comment", *Soviet Studies*, 14(1), p.75-80.
- Jasny Naum [1962b], "Soviet 'perspective' planning", in *Essays on the Soviet economy*, New-York, Praeger.
- Jasny Naum [1962b], "Soviet 'perspective' planning", in *Essays on the Soviet economy*, New-York, Praeger.
- Jasny Naum [1972], *Soviet economists of the twenties. Names to be remembered*, Cambridge University Press.
- Kantorovitch Leonid [1959], *Calcul économique et utilisation des ressources*, Paris, Dunod (1963).
- Kantorovitch Leonid [1959], *Calcul économique et utilisation des ressources*, Paris, Dunod (1963).
- Kaser Michael [1990], "The technology of decontrol: some macroeconomic issues", *The Economic Journal*, 100(401), juin, p.596-615.
- Lange Oskar [1969], *Theory of reproduction and accumulation*, Oxford, New-York, Pergamon Press, traduit du polonais par Jösef Stadler, traduction dirigée par P.F Knightsfield, *Teoria reprodukci i akumulacji* (1961) (1965).
- Lange Oskar [1969], *Theory of reproduction and accumulation*, Oxford, New-York, Pergamon Press, traduit du polonais par Jösef Stadler, traduction dirigée par P.F Knightsfield, *Teoria reprodukci i akumulacji* (1961) (1965).
- Leontief Estelle [1987], *Genia and Wassily, A Russian-American memoir*, Somerville (Mass), Zephyr Press.



Leontief Wassily [1925], “The balance of the economy of the USSR, A methodological analysis of the work of the Central Statistical Administration”, in Spulber [1964b], p.88-9, traduction du russe, “Balans narodnogo khoziaistva SSSR”, *Planovoe khoziaistvo*, 12, p.254- 258; aussi in Leontief Wassily, *Essays in Economics*, 2 ; Première publication en allemand, “Die bilanz der Russischen Volkswirtschaft. Eine methodologische Untersuchung”, *Weltwirtschaftliches Archiv*, 22(2), Octobre, p.338-344; et in *Weltwirtschaftliches Archiv. Chronik und Archivalien*, 22 (II), p.265-269.

Leontief Wassily [1925], “The balance of the economy of the USSR, A methodological analysis of the work of the Central Statistical Administration”, in Spulber [1964b], p.88-94, traduction du russe, “Balans narodnogo khoziaistva SSSR”, *Planovoe khoziaistvo*, 12, p.254- 258; aussi in Leontief Wassily, *Essays in Economics*, 2 ; Première publication en allemand, “Die bilanz der Russischen Volkswirtschaft. Eine methodologische Untersuchung”,

Leontief Wassily [1927], « Über die Theory und Statistik der Konzentration », *Jahrbücher für Nationalökonomie und Statistik*, 126, mars, traduction anglaise, Leontief Wassily [1985], “The theory and statistical description of concentration”, in *Essays in economics: theories, theorizing, facts and policies*, Transaction publishers, p.258-271.

Leontief Wassily [1927], « Über die Theory und Statistik der Konzentration », *Jahrbücher für Nationalökonomie und Statistik*, 126, mars, traduction anglaise, Leontief Wassily [1985], “The theory and statistical description of concentration”, in *Essays in economics: theories, theorizing, facts and policies*, Transaction publishers, p.258

Leontief Wassily [1928], “Die Wirtschaft als Kreislauf”, *Archiv für Sozialwissenschaft und Sozialpolitik*, 60, p.577-623, traduction anglaise “The economy as a circular flow”, *Structural change and economic dynamics*, (1990), 2, p.177-212.

Leontief Wassily [1928], “Die Wirtschaft als Kreislauf”, *Archiv für Sozialwissenschaft und Sozialpolitik*, 60, p.577-623, traduction anglaise “The economy as a circular flow”, *Structural change and economic dynamics*, (1990), 2, p.177-212.

Leontief Wassily [1936], “Quantitative input and output relations in the economic systems of the United States”, *The Review of Economic Statistics*, 18(3), August, p. 105-125.

Leontief Wassily [1937], “Interrelations of prices, output, savings and investment”, *The Review of Economic Statistics*, 19(3), August, p. 109-132.

Leontief Wassily [1941], *The Structure of American Economy*, Cambridge, Harvard University Press.

Levine Hebert S. [1964], “The Russian economic ‘balance’ and input-output analysis: a reply”, *Soviet Studies*, 15(3), p.352-356.

Levine Herbert S. [1962], “Input-output analysis and soviet planning”, *The American Economic Review*, 52(2), Papers and Proceedings, p.127-137.

Tretyakova Albina et Birman Igor [1976], “Input-output analysis in the USSR”, *Soviet Studies*, 28(2), p.157-186.

Vanoli André [2002], *Histoire de la comptabilité nationale*, Paris, La Découverte.

*Weltwirtschaftliches Archiv*, 22(2), Octobre, p.338-344; et in *Weltwirtschaftliches Archiv. Chronik und Archivalien*, 22 (II), p.265-269.

# ANNEXES

## ENCADRE1 : DISTRIBUTION DU RAPPORT DE DEUX VARIABLES

### Distribution du rapport de deux variables indépendantes

Soient deux variables indépendantes X et Y de densité de probabilité respective  $f_X$  et  $f_Y$ . Soit le rapport de ces deux variables  $Z = \frac{X}{Y}$  alors :

$$P(Z \leq z) = P\left(\frac{X}{Y} \leq z\right) = \int_{+\infty}^{-\infty} f_Y(y) P\left(\frac{X}{Y} \leq z\right) dy \quad (1)$$

$$\begin{aligned} & \int_0^{+\infty} f_Y(y) P(X \leq zy) dy + \int_{-\infty}^0 f_Y(y) P(X \geq zy) dy \\ &= \int_0^{+\infty} f_Y(y) \int_{-\infty}^{zy} f_X(x) dx dy + \int_{-\infty}^0 f_Y(y) \int_{zy}^{+\infty} f_X(x) dx dy \end{aligned} \quad (2)$$

$$= \int_0^{+\infty} f_Y(y) \int_{-\infty}^z f_X(xy) y dx dy - \int_{-\infty}^0 f_Y(y) \int_{-\infty}^z f_X(xy) y dx dy \quad (3)$$

Une dérivation fournit la densité de probabilité de présence du rapport.

$$f_Z(z) = \int_0^{+\infty} f_Y(y) f_X(zy) y dy - \int_{-\infty}^0 f_Y(y) f_X(zy) y dy \quad (4)$$

Cette formule se réduit à 2 fois son premier terme si les deux variables sont des distributions paires.

### Rapport des lois normales

- a- dans le cas du rapport de deux lois normales centrées réduites, le rapport suit une distribution de Breit-Wigner (ou de Cauchy).

$$f_Z(z) = \frac{1}{\pi} \times \frac{1}{1+z^2} \quad (5)$$

Qui en fait suit une loi de student à un degré de liberté.

Dans le cas de deux lois normales centrées, d'écart types respectifs  $\sigma_X$  et  $\sigma_Y$ , le rapport qui suit :

$$f_Z(z) = \frac{1}{\pi} \times \frac{\sigma_X}{\sigma_Y} \frac{1}{1 + \frac{\sigma_Y^2}{\sigma_X^2} z^2} \quad (6)$$

Cette distribution ne possède pas de moment d'ordre supérieur à la moyenne. Dans les cas général, deux lois normales centrées respectivement en  $m_i$  et d'écart type  $\sigma_i$  le rapport qui suit :

$$f_Z(z) = \frac{1}{\pi} \frac{\sigma_y}{\sigma_x} \times \frac{1}{1 + \frac{\sigma_y^2}{\sigma_x^2} z^2} e^{-\frac{1}{2} \left[ \frac{m_x^2 + m_y^2}{\sigma_x^2 + \sigma_y^2} \right]} \times \left[ 1 + \alpha(z) e^{\frac{1}{2} \alpha^2(z)} \int_0^{\alpha(z)} e^{-\frac{u^2}{2}} du \right]$$

$$\text{Avec } \alpha(z) = \frac{\frac{m_y}{\sigma_y^2} + \frac{m_x}{\sigma_x^2} z}{\sqrt{\frac{1}{\sigma_x^2} + \frac{1}{\sigma_y^2}}} = \frac{m_x}{\sigma_y} \times \frac{\frac{m_y}{m_x} + \frac{\sigma_y^2}{\sigma_x^2} z}{\sqrt{1 + \frac{\sigma_y^2}{\sigma_x^2} z^2}}$$

Pour un écart type du dénominateur grand devant sa moyenne, cette distribution se rapproche de la Breit-Wigner décrite à l'équation 5. Cependant, pour un écart type de dénominateur faible devant sa moyenne

( $(m_y \gg \sigma_y)$ ), soit en pratique  $m_y \gg 3\sigma_y$ . R.C. GEARY (1970) a prouvé que la variable  $\frac{m_y z - m_x}{\sigma_y^2 z^2 + \sigma_x^2}$

suivait approximativement une loi normale centrée réduite.

La moyenne de z est alors  $m = \frac{m_1}{m_2}$ .

En admettant, aux dépens de la rigueur mathématique que z ne s'en écarte pas trop, alors il s'ensuit que z suit une gaussienne centrée en m et d'écart type.

$$\frac{\sigma^2}{m^2} = \frac{\sigma_x^2}{m_x^2} + \frac{\sigma_y^2}{m_y^2}$$

## ENCADRE 2 : LES JUSTIFICATIFS DU CHOIX DE L'ESTIMATEUR

Soit  $B = \frac{t_{y_1}}{t_{y_2}}$  un ratio de deux quantités  $t_{y_1}$  et  $t_{y_2}$

L'estimateur de B est obtenu en estimant séparément le numérateur et le dénominateur. Cet estimateur  $\hat{B}$  est un ratio de deux variables aléatoires,  $\hat{t}_{y_1}$  et  $\hat{t}_{y_2}$ .  $\hat{B}$  est une fonction non linéaire de variables aléatoires.

Déterminer une expression exacte de son biais ou de sa variance est donc virtuellement impossible.

Pour contourner cette difficulté, il convient d'approximer la fonction non linéaire par une fonction linéaire des observations.

A cet effet, on linéarise la fonction non linéaire, en utilisant la linéarisation par séries de Taylor.

Linéarisation par série de Taylor du 1<sup>er</sup> ordre

Soit  $\theta = f(t_1, \dots, t_q)$ . On acceptera que  $\theta$  peut être approché

$$\text{par : } \hat{\theta} = f(\hat{t}_1, \dots, \hat{t}_q) = f(t_1, \dots, t_q) + \sum_{j=1}^q \left. \frac{\partial f(\hat{t}_1, \dots, \hat{t}_q)}{\partial \hat{t}_j} \right|_{\hat{t}_1=t_1, \dots, \hat{t}_q=t_q} (\hat{t}_j - t_j)$$

Soit  $B = f(t_{y_1}, t_{y_2})$ , un estimateur de B est  $\hat{B} = f(\hat{t}_{y_1}, \hat{t}_{y_2}) = \frac{\hat{t}_{y_1}}{\hat{t}_{y_2}}$

dans le cas d'espèce  $B = \frac{t_{y_1}}{t_{y_2}}$  et  $\hat{B} = \frac{\hat{t}_{y_1}}{\hat{t}_{y_2}}$

$$\hat{B} = f(\hat{t}_{y_1}, \hat{t}_{y_2}), \text{ on évalue les dérivées partielles } \left. \frac{\partial f(\hat{t}_{y_1}, \hat{t}_{y_2})}{\partial \hat{t}_{y_1}} \right|_{\hat{t}_{y_1}=t_{y_1}, \dots, \hat{t}_{y_2}=t_{y_2}} \text{ et } \left. \frac{\partial f(\hat{t}_{y_1}, \hat{t}_{y_2})}{\partial \hat{t}_{y_2}} \right|_{\hat{t}_{y_1}=t_{y_1}, \dots, \hat{t}_{y_2}=t_{y_2}} \text{ on a } \left. \frac{\partial f(\hat{t}_{y_1}, \hat{t}_{y_2})}{\partial \hat{t}_{y_1}} \right|_{\hat{t}_{y_1}=t_{y_1}, \dots, \hat{t}_{y_2}=t_{y_2}} = \frac{1}{t_{y_2}}$$

$$\left. \frac{\partial f(\hat{t}_{y_1}, \hat{t}_{y_2})}{\partial \hat{t}_{y_2}} \right|_{\hat{t}_{y_1}=t_{y_1}, \dots, \hat{t}_{y_2}=t_{y_2}} = \frac{-t_{y_1}}{t_{y_2}^2}$$

Ce qui donne

$$\hat{B} = f(\hat{t}_{y_1}, \hat{t}_{y_2}) = f(t_{y_1}, t_{y_2}) + \frac{1}{t_{y_2}} (\hat{t}_{y_1} - t_{y_1}) - \frac{t_{y_1}}{t_{y_2}^2} (\hat{t}_{y_2} - t_{y_2}) + \text{reste} \approx B + \frac{1}{t_{y_2}} (\hat{t}_{y_1} - t_{y_1}) - \frac{t_{y_1}}{t_{y_2}^2} (\hat{t}_{y_2} - t_{y_2})$$

$$\hat{B} - B = \frac{1}{t_{y_2}} \left[ (\hat{t}_{y_1} - t_{y_1}) - \frac{t_{y_1}}{t_{y_2}} (\hat{t}_{y_2} - t_{y_2}) \right]$$

$$\hat{B} - B = \frac{1}{t_{y_2}} \left[ (\hat{t}_{y_1} - t_{y_1}) - \frac{t_{y_1}}{t_{y_2}} \hat{t}_{y_2} + t_{y_1} \right]$$

$$\hat{B} - B \approx \frac{1}{t_{y_2}} (\hat{t}_{y_1} - B \hat{t}_{y_2})$$

$$E(\hat{B} - B) \approx \frac{1}{t_{y_2}} [E(\hat{t}_{y_1}) - BE(\hat{t}_{y_2})] = \frac{1}{t_{y_2}} [t_{y_1} - B t_{y_2}] = 0$$

Un développement par séries de Taylor du 1<sup>er</sup> ordre montre donc que  $\hat{B}$  est approximativement sans biais de  $B$ .

Autrement dit, le biais de  $\hat{B}$  est considéré négligeable pour une taille d'échantillon grande.

### L'estimation de l'erreur de l'estimateur

Dans cette section, on cherche à trouver une borne supérieure à l'erreur effectuée sur l'estimateur.

Soit  $B = \frac{X}{Y}$  (1) le ratio de deux variables aléatoires  $X$  et  $Y$ .

Un estimateur naturel de  $B$  consiste à estimer séparément le numérateur et le dénominateur en (1). Un

estimateur de  $B$  est donné par  $\hat{B} = \frac{\hat{X}}{\hat{Y}}$  (2).

Cependant, la question qui se pose est de savoir si  $\hat{B}$  est un estimateur sans biais de  $B$ .

$$E(\hat{B}) = E\left(\frac{\hat{X}}{\hat{Y}}\right) \neq \frac{E(\hat{X})}{E(\hat{Y})}$$

Autrement dit, l'espérance d'un ratio n'est pas égale au ratio des espérances. Peut-on alors, évaluer le biais de  $\hat{B}$ ? Il n'est pas facile d'obtenir l'expression exacte du biais de  $\hat{B}$ . On se contentera d'obtenir une borne supérieure du biais de  $\hat{B}$ .

Soit  $\varepsilon(X)$  et  $\varepsilon(Y)$  les erreurs faites respectivement sur l'estimateur de  $X$  et de  $Y$ .

On a l'égalité  $B = \frac{\hat{X} + \varepsilon(X)}{\hat{Y} + \varepsilon(Y)}$

$$B = \frac{\hat{X} + \varepsilon(X)}{\hat{Y} + \varepsilon(Y)} \text{ ce qui implique } B = \frac{\hat{X}}{\hat{Y} + \varepsilon(Y)} + \frac{\varepsilon(X)}{\hat{Y} + \varepsilon(Y)}$$

Le principal étant de trouver une borne supérieure à l'erreur, on suit les étapes suivantes.

$$\hat{Y} + |\varepsilon(Y)| \geq |\varepsilon(Y)| \Rightarrow \frac{|\varepsilon(X)|}{\hat{Y} + |\varepsilon(Y)|} \leq \frac{|\varepsilon(X)|}{|\varepsilon(Y)|}$$

Donc une borne supérieure de l'erreur devient  $\frac{|\varepsilon(X)|}{|\varepsilon(Y)|}$ .

En d'autres termes, la borne supérieure de l'erreur faite sur l'estimateur est :  $\frac{|\varepsilon(X)|}{|\varepsilon(Y)|}$

Et le ratio  $B$  peut se réécrire sous la forme  $B \approx \frac{\hat{X}}{\hat{Y}} + \frac{|\varepsilon(X)|}{|\varepsilon(Y)|}$

### ENCADRE 3 : TEST DE NORMALITE

#### Tests de Normalité

Le **test de Jarque-Bera** cherche à déterminer si des données suivent une loi normale. On a :

H0: les données suivent une loi normale

H1: les données ne suivent pas une loi normale

$$JB = \frac{n - k}{6} \left( S^2 + \frac{(K - 3)^2}{4} \right)$$

La quantité JB suit asymptotiquement une loi du  $\chi^2$  à 2 degrés de liberté, le test s'effectuant sur les résidus.

On teste donc l'hypothèse de normalité des résidus au seuil  $\alpha$ .

Avec:

n = Nombre d'observations

k = Nombre de variables explicatives

S = Coefficient d'asymétrie : Moment d'ordre 3 d'une variable centrée-réduite

K = Kurtosis : Moment d'ordre 4 d'une variable centrée-réduite

Comme estimateur du moment d'ordre 1, on prend la moyenne.

Une loi normale a un coefficient d'asymétrie = 0 et une kurtosis = 3. On saisit alors que si les données suivent une loi normale, le test vaut alors 0 et on accepte (ne rejette pas) Ho au seuil  $\alpha$ .

#### JARQUE-BERA : test de normalité

- $H_0$  : normalité.
- $H_1$  : écart à la normalité.
- statistique de test :  $JB = \beta_1 \left( \frac{n}{6} \right) + (\beta_2 - 3)^2 \left( \frac{n}{24} \right)$  où  $\beta_1$  et  $\beta_2$  correspondent aux *kurtosis* et au *skewness*.
- rejet de  $H_0$  :  $JB > \chi_2^2(1 - \alpha)$

## ENCADRE 4 : TEST DE STATIONNARITE

Le fait qu'un processus soit stationnaire ou non conditionne le choix de la modélisation que l'on doit adopter. En règle générale, si l'on s'en tient notamment à la méthodologie de Box et Jenkins, si la série étudiée est issue d'un processus stationnaire, on cherche alors le meilleur modèle parmi la classe des processus stationnaire pour la représenter, puis on estime ce modèle. En revanche si la série est issue d'un processus non stationnaire, on doit avant toutes choses, chercher à la "stationnariser", c'est à dire trouver une transformation stationnaire de ce processus. Puis, on modélise et l'on estime les paramètres associés à la composante stationnaire.

La difficulté réside dans le fait qu'il existe différentes sources de non stationnarité et qu'à chaque origine de la non stationnarité est associée une méthode propre de stationnarisation.

Nous rappelons au passage qu'il existe deux classes de processus non stationnaires, selon la terminologie de Nelson et Plosser (1982) : les processus TS (Time Stationary) et les processus DS (Differency Stationary). Chacune de ces classes de processus présentant des méthodes de stationnarité propres.

Mais au delà des enjeux de modélisation économétriques, l'origine de la non stationnarité a de très fortes implications sur l'analyse économique des séries que l'on étudie. En particulier pour les processus DS il existe une propriété de persistance des chocs qui n'existe pas dans les processus TS: Une telle hypothèse implique par exemple que si les séries macroéconomiques satisfont une représentation de type DS; l'impact des chocs conjoncturels peut avoir un effet permanent sur le niveau de la série étudiée.

### Test de racine unitaire

La discussion suivante souligne les grandes lignes des tests de racine unitaire. Pour des raisons de commodité, la discussion sera brève. Les utilisateurs qui voudraient avoir le détail, devraient consulter les sources originales et les références standards (cf. par exemple, Davidson et MacKinnon, 1993, Chapitre 20, Hamilton, 1994, Chapitre 17, et Hayashi, 2000, Chapitre 9).

Considérons un processus AR(1) simple:

$$(1) \quad y_t = \rho y_{t-1} + x_t' \delta + \varepsilon_t$$

où  $x_t$  est un régresseur exogène facultatif qui est composé d'une constante, ou d'une constante avec tendance.  $\rho$  et  $\delta$  sont des paramètres qui peuvent être estimés et  $\varepsilon_t$  est supposé être un bruit blanc.

Si  $|\rho| \geq 1$ ,  $y$  est une série non stationnaire et la variance augmente avec le temps, et tend vers l'infini. Si  $|\rho| < 1$ ,  $y$  est une série stationnaire avec tendance. Donc, cette hypothèse de stationnarité avec tendance peut être évaluée en testant ou non si la valeur absolue de  $\rho$  est strictement inférieure à un.



## Test de Dickey-Fuller Augmenté (ADF)

Le test de Dickey-Fuller standard est effectué en estimant l'équation (1) après avoir retrancher  $y_{t-1}$  des deux membres de l'équation.

$$(2) \Delta y_t = \alpha y_{t-1} + x_t' \delta + \varepsilon_t$$

où  $\alpha = \rho - 1$ . L'hypothèse nulle et alternative peut être écrit comme suit :

$$H_0: \alpha = 0$$

$$(3) H_1: \alpha < 0$$

Et évalué en utilisant le taux conventionnel pour :

$$(4) t_\alpha = \hat{\alpha} / (se(\hat{\alpha}))$$

où  $\hat{\alpha}$  est l'estimation de  $\alpha$ , et  $se(\hat{\alpha})$  est le coefficient de l'erreur standard.

Dickey et Fuller (1979) ont montré que sous l'hypothèse nulle de racine unitaire, cette statistique ne suit pas la distribution conventionnelle du t de student, et ils ont trouvé des résultats asymptotiques et des valeurs critiques simulées pour des tests variés et la taille de l'échantillon. Plus récemment, MacKinnon (1991, 1996) a réalisé un ensemble de simulations plus élargi que celles présentées sous forme de tableaux par Dickey et Fuller. De plus, MacKinnon a effectué des travaux de simulations dont les résultats permettent le calcul des valeurs critiques de Dickey-Fuller.

Le test de racine unitaire simple de Dickey-Fuller décrit au-dessus est valide si seulement les séries suivent des processus AR(1). Si les séries sont corrélées avec un nombre élevé de retards, l'hypothèse de perturbation du bruit blanc  $\varepsilon_t$  est violée. Le test de Dickey-Fuller augmenté (ADF) construit une correction paramétrique d'un haut niveau de corrélation en supposant que la série Y suit un processus AR(p) et en ajoutant p retard en différence de la variable dépendante y au membre de droite de l'équation du test de régression:

$$(5) \Delta y_t = \alpha y_{t-1} + x_t' \delta + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \dots + \beta_p \Delta y_{t-p} + v_t$$

Cette spécification ajoutée est donc utilisée pour tester (3) utilisant le rapport (4). Un important résultat obtenu par Fuller est que la distribution asymptotique du rapport est indépendante du nombre de retard en différence première inclus dans la régression du test de Dickey-Fuller augmenté. En outre, alors que l'hypothèse qui suit un processus autorégressif (AR) peut sembler restrictive, Said et Dickey (1984) ont démontré que le test ADF est asymptotiquement valide en présence des composantes de moyenne mobile, à condition qu'un nombre suffisant de retard en différence soit inclus dans le test de régression.

Il se pose cependant des problèmes de choix lors de la mise en œuvre du test ADF.

En effet, il faut choisir d'inclure ou non des variables exogènes dans le test de régression. Le choix d'inclure dans le test de régression une constante ou une constante avec tendance temporelle linéaire, ou aucun des deux. L'approche généralement retenue sera de faire le test avec une constante et un trend linéaire, dans la mesure où les deux autres cas sont justes des cas spécifiques de cette spécification générale. Toutefois, inclure des régresseurs non pertinents dans la régression va réduire la force du test de rejeter la nullité de la racine unitaire. La recommandation standard est de choisir une spécification qui est une description plausible des données sous les hypothèses nulles et alternatives (cf. Hamilton (1994a, p. 501)).

## **ENCADRE 5 : L'ESTIMATEUR COMBINAISON LINEAIRE DE LA MOYENNE ET DE LA MEDIANE**

Une statistique de test a été proposée par Turkheimer et al. dans le cadre des analyses d'images médicales RMN. On reste sur l'idée intuitive d'une statistique de test proportionnelle à la différence entre les moyennes et réduite par un facteur d'échelle.

Classiquement on peut choisir la moyenne empirique de l'échantillon,  $\bar{\mu}$  ou la médiane empirique de l'échantillon  $\hat{\mu}$ .

La médiane présente l'avantage d'être plus robuste en cas de valeurs extrêmes, c'est-à-dire quand on a des distributions à queues plus lourdes qu'une distribution normale. En effet, contrairement à la moyenne empirique, la médiane est un estimateur robuste de la tendance centrale d'une distribution. Elle n'est pas affectée par des valeurs très petites ou très grandes qui peuvent par contre biaiser le calcul de la moyenne empirique.

Moyenne empirique et médiane empirique sont toutes deux des caractéristiques de tendance centrale. Choisir la moyenne, c'est choisir de minimiser l'écart quadratique, choisir la médiane, c'est minimiser l'écart absolu moyen.

$$\bar{\mu} = \arg \min_x \text{EQM}(x) = \arg \min_x \sum_{j=1}^n (X_{ij} - x)^2$$

$$\hat{\mu} = \arg \min_x \text{EAM}(x) = \arg \min_x \sum_{j=1}^n |X_{ij} - x|$$

Plutôt que de faire un choix entre médiane et moyenne, Turkheimer et al., utilisent une approche assez intuitive initialement développée par Samuel-Cahn en 1994: une combinaison linéaire de ces deux estimateurs.

$$\left(\hat{\mu}\right)_\lambda = \lambda \bar{\mu} + (1 - \lambda) \hat{\mu} \quad \text{avec } \lambda > 0$$

Où  $\lambda$  est un facteur de pondération.

On appellera cet estimateur  $(\text{LCMM}(X))_\lambda$ , c'est-à-dire, en anglais, la Combinaison

Linéaire de la Moyenne et de la Médiane. Cet estimateur présente de nombreux avantages. En faisant varier la valeur de  $\lambda$ , il peut tirer profit de la bonne efficacité de la moyenne empirique quand la distribution est proche d'une distribution normale et cependant bénéficier de la robustesse de la médiane quand les queues de distribution sont plus lourdes et/ou qu'il y a des valeurs extrêmes.

Le facteur de pondération  $\lambda$  peut prendre n'importe quelle valeur positive. Dans le cas d'une distribution avec des queues plus légères qu'une gaussienne, la valeur optimale de  $\lambda$  sera supérieure à 1; par exemple, le mélange optimal entre la moyenne et la médiane pour estimer l'espérance d'une distribution uniforme à un facteur de pondération égal à 3/2.

Quand on s'attend à avoir des données dont la loi a des queues au moins aussi lourdes qu'une loi normale, les valeurs optimales de  $\lambda$  sont situées dans l'intervalle 0 et 1.

Pour calculer la statistique de test, nous avons aussi besoin d'un estimateur d'échelle  $s(\mathbf{X}^1, \mathbf{X}^2)$ . Il existe une large variété d'estimateurs robustes de l'écart-type; en raison de sa robustesse dans le cadre des échantillons de petite taille, on peut choisir un estimateur nommé l'écart-type empirique agrégé proposé par D'Agostino et Cureton en 1973. La statistique de test  $Z$  a alors la forme finale:

$$Z_\lambda = \frac{(\text{LCMM}(\mathbf{X}^1))_\lambda - (\text{LCMM}(\mathbf{X}^2))_\lambda}{s(\mathbf{X}^1, \mathbf{X}^2)}$$

La dernière étape pour être en mesure de calculer la statistique de test est de déterminer la valeur optimale de  $\lambda$  en fonction des données. On utilise une technique appelée MVA (Variance Minimum Adaptive) (Turkheimer et al.) qui appartient à la classe des techniques d'estimation adaptatives par « bootstrap » originellement introduites par Léger et Romano en 1990. Il a été montré que la technique MVA avait de bonnes propriétés pour des échantillons petits pour une utilisation avec divers estimateurs robustes.

**Tableau 3 : Test de normalité sur les branches**

VARIABLES	RMSE			THEIL			Mean	Median	Std. Dev.	Jarque-Bera	Probability
	HW Additive	HW multiplicative	HW no seasonal	HW Additive	HW multiplicative	HW no seasonal					
VAL10	31067,78	<b>30086,10</b>	38489,37	0,060	<b>0,058</b>	0,076	162402,8	144051,50	61963,19	0,847706	0,6545
VAL20	42776,56	<b>40353,73</b>	52809,58	0,210	<b>0,199</b>	0,257	97747,44	93897,78	33744,18	1,702693	0,4268
VAL30	22403,68	21488,51	<b>20616,00</b>	0,063	0,061	<b>0,059</b>	116806,1	97874,78	55692,27	2,000308	0,3678
VAL40	4530,61	<b>4347,95</b>	4310,17	0,066	<b>0,064</b>	0,064	22638,53	13243,58	13971,92	2,702178	0,2590
VAL50	19102,78	<b>18463,68</b>	19518,65	0,082	<b>0,079</b>	0,085	77553,5	59208,22	40660,64	2,09798	0,3503
VAL60	<b>7583,94</b>	8064,70	7870,42	<b>0,022</b>	0,024	0,023	48172,63	34335,30	23856,95	3,038718	0,2189
VAL70	21038,11	<b>20997,34</b>	21727,66	0,042	<b>0,042</b>	0,043	163927,1	137905,50	72389,10	2,315631	0,3142
VAL80	<b>45979,48</b>	49348,88	49568,67	<b>0,138</b>	0,149	0,151	97094,11	89958,85	32662,64	0,256021	0,8798
VAL90	<b>11473,24</b>	12042,14	14203,34	<b>0,048</b>	0,050	0,060	79014,23	68205,04	36665,57	1,388445	0,4995
VAL100	13826,57	13227,38	<b>11485,31</b>	0,062	0,059	<b>0,052</b>	73477,78	56939,65	40151,40	2,238131	0,3266
VAL110	3466,95	<b>3450,56</b>	4083,14	0,029	<b>0,029</b>	0,035	33164,95	28940,68	9141,36	2,623337	0,2694
VAL120	8509,26	10459,22	<b>4778,14</b>	0,058	0,071	<b>0,033</b>	49583,09	34007,90	36504,81	2,975584	0,2259
VAL130	2911,65	3157,62	<b>2108,48</b>	0,055	0,059	<b>0,040</b>	17382,38	12138,50	10856,35	3,098764	0,2124
VAL140	5173,71	<b>4001,74</b>	4987,92	0,089	<b>0,069</b>	0,087	18575,06	13750,99	10942,79	8,624958	0,0134
VAL150	11726,27	<b>10443,11</b>	10659,17	0,050	<b>0,044</b>	0,045	76119,68	61142,31	36118,82	2,270938	0,3213
VAL160	1739,08	1520,41	<b>1103,95</b>	0,093	0,081	<b>0,060</b>	6125,424	3451,43	4777,07	4,242226	0,1199
VAL170	2727,58	3034,37	<b>2497,40</b>	0,037	0,041	<b>0,034</b>	25071,56	19089,05	14468,66	2,397619	0,3016
VAL180	4674,80	4925,07	<b>4201,95</b>	0,050	0,052	<b>0,045</b>	31967,35	24464,20	19289,26	2,273479	0,3209
VAL190	34558,90	36921,54	<b>35454,14</b>	0,117	0,125	<b>0,122</b>	86113,7	59076,90	67737,55	6,929745	0,0313
VAL200	24187,04	24384,99	<b>17161,24</b>	0,091	0,092	<b>0,065</b>	89795,65	56706,32	68641,87	2,732941	0,2550
VAL 210	3128,27	3283,69	<b>2348,43</b>	0,048	0,051	<b>0,036</b>	21507,89	15202,86	13493,84	3,557553	0,1688
VAL 220	5907,51	6500,92	<b>4803,86</b>	0,055	0,060	<b>0,045</b>	34916,63	19693,15	29595,81	6,304221	0,0428
VAL 230	6140,36	5850,55	<b>4964,91</b>	0,067	0,064	<b>0,055</b>	31132,8	21526,70	20115,16	2,747325	0,2532
VAL 240	1282,50	1584,60	<b>1109,02</b>	0,060	0,074	<b>0,052</b>	7171,464	4340,35	5337,50	3,444973	0,1786
VAL 260	1708,08	6187,29	<b>1652,61</b>	0,186	0,671	<b>0,178</b>	2890,641	2130,82	1963,69	2,87655	0,2373
VAL 270	4128,37	3805,37	<b>3050,10</b>	0,035	0,033	<b>0,026</b>	38325,53	26980,87	21319,75	2,97252	0,2262
VAL 280	12230,95	11867,44	<b>10619,29</b>	0,042	0,041	<b>0,037</b>	94649,16	76818,16	50122,51	3,148878	0,2071
VAL 290	38656,89	35813,32	<b>33227,58</b>	0,052	0,048	<b>0,045</b>	250876,6	179208,80	181176,40	3,573515	0,1675
VAL 300	<b>47977,56</b>	<b>47971,90</b>	55946,07	<b>0,030</b>	<b>0,030</b>	0,036	512048,4	375379,80	249584,20	2,515983	0,2842
VAL 310	9410,67	7373,22	<b>6778,76</b>	0,050	0,040	<b>0,037</b>	61991,97	44503,44	35650,65	2,690185	0,2605
VAL 320	11409,91	<b>10897,73</b>	11902,44	0,038	<b>0,036</b>	0,040	99298,4	91073,65	44525,92	1,325792	0,5154
VAL 330	21185,85	<b>20430,82</b>	20640,13	0,036	<b>0,035</b>	0,036	184633,7	169364,50	68708,83	0,893747	0,6396
VAL 340	18749,00	20052,67	<b>11617,32</b>	0,065	0,069	<b>0,041</b>	95683,56	64518,00	81192,28	8,431381	0,0148
VAL 350	12220,84	10187,21	<b>10047,30</b>	0,086	0,072	<b>0,071</b>	76269,13	57378,44	44430,97	2,404524	0,3005
VAL 360	13803,65	16136,89	<b>12413,54</b>	0,035	0,041	<b>0,032</b>	129064,5	92170,57	74871,73	2,778622	0,2492
VAL 370	18854,47	<b>15954,96</b>	16301,31	0,054	<b>0,046</b>	0,047	114297,7	76050,77	71249,11	3,425566	0,1804
VAL 380	21067,19	17751,99	<b>14538,77</b>	0,032	0,027	<b>0,022</b>	204014,4	166298,50	84059,15	2,863921	0,2388
VAL 390	11362,17	<b>10460,33</b>	10673,84	0,031	<b>0,019</b>	0,029	103303	100960,70	25060,76	1,406472	0,4950
VAL 400	4770,92	4992,88	<b>4216,74</b>	0,033	0,035	<b>0,029</b>	46303,24	34149,41	22549,39	2,701283	0,2591
VAL 410	8200,22	<b>6807,70</b>	7488,67	0,042	<b>0,035</b>	0,039	62918,58	48075,91	30433,53	2,485777	0,2886

**Tableau 4 : Liste des sigles**

<b>SIGLE</b>	<b>LIBELLE</b>
MAX	MAXIMUM
MIN	MINIMUM
BR	BRANCHE
RMSE	ROOT MEAN SQUARE ERROR
THEIL	COEFFICIENT DE THEIL
CI	CONSOMMATION INTERMEDIAIRE
PDB	PRODUCTION DISPONIBLE BRUTE
TEI	TABLEAU DES ECHANGES INTER INDUSTRIELS
FILTRE HP	FILTRE DE HODRICK ET PRESCOTT
HW	METHODE DE HOLT WINTERS
MCS	MATRICE DE COMPTABILITE SOCIALE
TRE	TABLEAU RESSOURCES EMPLOIS
TEI	TABLEAU DES COMPTES ECONOMIQUES INTEGRES
SCN	SYSTEME DE COMPTABILITE NATIONALE
TES	TABLEAU ENTRES SORTIES
TS	TREND STATIONARY PROCESSES
DS	DIFFERENCY STATIONNARY PROCESSES

**Tableau 5 : Calcul des erreurs quadratiques moyennes (RMSE) des produits disponibles par branches**

<b>Variabes</b>	<b>Moyenne estimée</b>	<b>Médiane estimée</b>	<b>RMSE_Br</b>
PDB 10	162 402,80	144 051,50	30 086,10
PDB 20	97 747,44	93 897,78	40 353,73
PDB 30	116 806,10	97 874,78	20 616,00
PDB 40	22 638,53	13 243,58	4 347,95
PDB 50	77 553,50	59 208,22	18 463,68
PDB 60	48 172,63	34 335,30	7 583,94
PDB 70	163 927,10	137 905,50	20 997,34
PDB 80	97 094,11	89 958,85	45 979,48
PDB 90	79 014,23	68 205,04	11 473,24
PDB 100	73 477,78	56 939,65	11 485,31
PDB 110	33 164,95	28 940,68	3 450,56
PDB 120	49 583,09	34 007,90	4 778,14
PDB 130	17 382,38	12 138,50	2 108,48
PDB 140	18 575,06	13 750,99	4 001,74
PDB 150	76 119,68	61 142,31	10 443,11
PDB 160	6 125,42	3 451,43	1 103,95
PDB 170	25 071,56	19 089,05	2 497,40
PDB 180	31 967,35	24 464,20	4 201,95
PDB 190	86 113,70	59 076,90	35 454,14
PDB 200	89 795,65	56 706,32	17 161,24
PDB 210	21 507,89	15 202,86	2 348,43
PDB 220	34 916,63	19 693,15	4 803,86
PDB 230	31 132,80	21 526,70	4 964,91
PDB 240	7 171,46	4 340,35	1 109,02
PDB 260	2 890,64	2 130,82	1 652,61
PDB 270	38 325,53	26 980,87	3 050,10
PDB 280	94 649,16	76 818,16	10 619,29
PDB 290	250 876,60	179 208,80	33 227,58
PDB 300	512 048,40	375 379,80	47 971,90
PDB 310	61 991,97	44 503,44	6 778,76
PDB 320	99 298,40	91 073,65	10 897,73
PDB 330	184 633,70	169 364,50	20 430,82
PDB 340	95 683,56	64 518,00	11 617,32
PDB 350	76 269,13	57 378,44	10 047,30
PDB 360	129 064,50	92 170,57	12 413,54
PDB 370	114 297,70	76 050,77	15 954,96
PDB 380	204 014,40	166 298,50	14 538,77
PDB 390	103 303,00	100 960,70	10 460,33
PDB 400	46 303,24	34 149,41	4 216,74
PDB 410	62 918,58	48 075,91	6 807,70

**Tableau 6 : Calcul des erreurs quadratiques moyennes des consommations intermédiaires par branches**

<b>Branches</b>	<b>Moyenne estimée</b>	<b>Médiane estimée</b>	<b>RMSE_CI</b>
PDB 10	22 225,79	21 017,80	832,71
PDB 20	23 275,97	22 600,14	1 770,93
PDB 30	21 300,64	20 126,55	644,41
PDB 40	7 763,79	6 297,86	260,76
PDB 50	33 927,24	29 327,39	1 661,44
PDB 60	24 528,72	20 712,85	730,66
PDB 70	72 576,13	65 620,22	2 798,89
PDB 80	90 488,08	94 657,08	7 008,85
PDB 90	67 387,80	64 422,77	1 920,93
PDB 100	63 576,13	55 986,64	1 529,01
PDB 110	18 700,41	16 719,06	511,06
PDB 120	40 565,70	31 368,17	614,88
PDB 130	11 116,29	8 884,69	239,77
PDB 140	15 444,46	12 004,63	631,30
PDB 150	43 886,59	39 523,06	1 099,16
PDB 160	2 971,00	2 117,38	80,63
PDB 170	14 929,34	13 743,16	362,13
PDB 180	22 384,99	19 892,74	495,26
PDB 190	82 335,56	46 617,65	5 220,46
PDB 200	58 850,60	48 299,69	1 742,28
PDB 210	14 988,79	11 724,34	416,88
PDB 220	23 540,16	14 003,50	659,52
PDB 230	18 263,16	15 021,33	504,85
PDB 240	5 626,41	4 106,92	139,60
PDB 260	1 871,80	1 767,14	181,43
PDB 270	15 565,68	12 306,49	245,76
PDB 280	44 145,02	31 787,43	1 143,10
PDB 290	181 999,20	133 625,80	3 736,49
PDB 300	146 721,60	128 206,70	3 365,47
PDB 310	47 300,68	39 861,20	776,61
PDB 320	77 801,36	72 680,13	1 836,43
PDB 330	95 702,73	91 278,27	2 167,84
PDB 340	35 032,81	22 158,25	1 218,15
PDB 350	22 514,22	18 778,95	492,09
PDB 360	9 304,37	7 573,63	181,77
PDB 370	39 731,19	34 863,34	1 352,50
PDB 380	41 100,38	30 090,29	987,00
PDB 390	16 024,08	15 534,84	626,74
PDB 400	10 689,96	7 215,65	308,76
PDB 410	21 201,30	18 469,28	430,66

**Tableau 7 : Coefficients Techniques de 2001 à 2004 et intervalles de confiance**

Estimation des coefficients techniques et intervalle de Confiance					Structure des Coefficients Techniques				
	Structure <sup>13</sup>	Erreur	Borne Inf	Borne Sup	Coef 2001	Coef 2002	Coef 2003	Coef 2004	Coef 2005
BR 10	13,67%	2,77%	10,92%	16,22%	13,24%	13,26%	13,30%	13,30%	13,13%
BR 20	24,04%	4,39%	19,73%	27,74%	23,20%	23,23%	23,27%	23,27%	23,08%
BR 30	18,64%	3,13%	16,68%	21,64%	17,83%	17,83%	17,84%	17,84%	17,80%
BR 40	35,19%	6,00%	29,37%	40,53%	34,99%	35,06%	35,16%	35,16%	34,65%
BR 50	44,34%	9,00%	40,25%	52,79%	45,24%	45,36%	45,54%	45,54%	44,64%
BR 60	51,44%	9,63%	45,85%	60,62%	47,50%	47,62%	47,82%	47,82%	46,85%
BR 70	44,04%	13,33%	31,10%	57,33%	46,23%	46,23%	49,29%	49,29%	57,53%
BR 80	91,67%	15,24%	84,59%	106,01%	94,67%	94,80%	94,99%	94,99%	94,01%
BR 90	86,59%	16,74%	75,03%	102,75%	84,46%	84,58%	84,75%	84,75%	83,89%
BR 100	89,15%	13,31%	78,97%	102,12%	84,53%	84,61%	84,74%	84,74%	84,09%
BR 110	56,50%	14,81%	41,76%	70,64%	55,74%	55,82%	55,94%	55,94%	55,33%
BR 120	83,15%	12,87%	72,73%	95,75%	80,87%	80,97%	81,11%	81,11%	80,40%
BR 130	66,75%	11,37%	53,57%	71,69%	60,04%	60,13%	60,25%	60,25%	59,62%
BR 140	84,67%	15,78%	68,93%	100,15%	80,51%	80,66%	80,87%	80,87%	79,81%
BR 150	58,45%	10,53%	52,32%	68,48%	58,83%	58,90%	59,01%	59,01%	58,48%
BR 160	50,95%	7,30%	42,69%	52,32%	45,63%	45,70%	45,82%	45,82%	45,24%
BR 170	62,00%	14,50%	45,89%	62,02%	53,89%	53,95%	54,06%	54,06%	53,54%
BR 180	72,55%	11,79%	59,98%	83,55%	66,01%	66,10%	66,24%	66,24%	65,55%
BR 190	99,98%	14,72%	87,16%	97,60%	92,51%	92,54%	92,59%	92,59%	92,35%
BR 200	70,20%	10,15%	57,67%	77,97%	62,21%	62,33%	62,51%	62,51%	61,63%
BR 210	70,63%	17,75%	55,14%	88,14%	66,14%	66,29%	66,51%	66,51%	65,41%
BR 220	69,66%	13,73%	59,77%	82,96%	67,56%	67,75%	68,04%	68,04%	66,59%
BR 230	61,24%	10,17%	56,35%	70,82%	58,49%	58,59%	58,73%	58,73%	58,02%
BR 240	80,92%	12,59%	76,35%	92,61%	77,46%	77,60%	77,80%	77,80%	76,80%
BR 260	58,72%	10,98%	48,05%	70,01%	67,43%	67,59%	67,84%	67,84%	66,60%
BR 270	41,50%	8,06%	36,74%	49,18%	41,72%	41,77%	41,84%	41,84%	41,48%
BR 280	46,88%	10,76%	36,74%	58,27%	54,91%	55,17%	55,56%	55,56%	53,59%
BR 290	75,15%	11,25%	70,81%	85,62%	73,63%	73,74%	73,90%	73,90%	73,09%
BR 300	28,98%	7,01%	24,90%	35,66%	28,96%	28,99%	29,04%	29,04%	28,79%
BR 310	78,76%	11,46%	66,46%	89,37%	73,51%	73,64%	73,84%	73,84%	72,84%
BR 320	78,81%	16,85%	63,41%	95,50%	77,48%	77,57%	77,72%	77,72%	77,00%
BR 330	52,60%	10,61%	43,57%	63,03%	50,37%	50,59%	50,94%	50,94%	49,23%
BR 340	37,69%	10,49%	27,30%	47,26%	36,15%	38,68%	36,43%	36,43%	35,69%
BR 350	30,41%	4,90%	26,97%	35,14%	30,23%	30,27%	30,33%	30,33%	30,03%
BR 360	7,37%	1,46%	6,65%	8,75%	7,25%	7,26%	7,28%	7,28%	7,21%
BR 370	36,06%	8,48%	26,71%	43,66%	29,18%	29,20%	29,23%	29,23%	29,09%
BR 380	20,16%	6,79%	13,60%	27,18%	23,52%	24,09%	26,26%	26,26%	24,69%
BR 390	15,59%	5,99%	9,62%	21,35%	10,26%	10,07%	11,08%	11,08%	14,71%
BR 400	23,12%	7,32%	16,08%	30,72%	27,06%	29,38%	30,82%	30,82%	44,02%
BR 410	33,95%	0,06	29,65%	40,06%	33,20%	33,25%	33,32%	33,32%	32,96%

$$\text{Erreur : } \left[ \frac{RMSE(\hat{CI})}{RMSE(\hat{P\hat{D}B\text{Branche}})} \right]$$

Borne Inférieure/Borne Supérieure : confert page 7

<sup>13</sup> Cette structure représente une combinaison linéaire de la moyenne et de la médiane, avec un poids de  $\lambda = 0,2$  pour la moyenne et  $(1 - \lambda) = 0,8$  pour la médiane. Ce poids est obtenu par minimisation de la variance de l'erreur.



**Tableau 8 : Structure médiane de la part des produits dans la consommation intermédiaire**

<b>Libellé des Produits</b>	<b>Structure Médiane</b>
010 agriculture vivrière	5,19%
020 agriculture industrielle ou d'exportation	3,38%
030 élevage et chasse	3,53%
040 sylviculture, exploitation forestière,	0,48%
050 pêche	2,13%
060 activités extractives	1,24%
070 Transformation et conservation de viande, poisson	4,97%
080 fabrication de corps gras alimentaires	3,24%
090 travail de grains, fabrication de produits	2,46%
100 Fabrication de produits alimentaires céréaliers	2,05%
110 fabrication de sucre, transformation	1,04%
120 fabrication de produits alimentaires n.c.a	1,23%
130 fabrication de boissons	0,44%
140 fabrication de produits à base de tabac	0,50%
150 égrenage de coton et fabrication des textiles	2,20%
160 fabrication du cuir; fabrication	0,12%
170 travail du bois et fabrication d'articles	0,69%
180 fabrication de papier, carton,	0,88%
190 raffinage pétrole, cokéfaction,	2,13%
200 fabrication de produits chimiques	2,04%
210 fabrication de produits en caoutchouc	0,55%
220 fabrication de verre, poterie	0,71%
230 métallurgie, fonderie, fabrication	0,78%
240 fabrication de machines	0,16%
250 fabrication d'équipements, d'appareils	0,00%
260 construction de matériels de transports	0,08%
270 fabrication de mobilier, fabrication n.c.a.,	0,97%
280 électricité, gaz et eau	2,77%
290 construction	6,46%
300 commerce	13,53%
310 services de la réparation	1,60%
320 services d'hébergement et restauration	3,28%
330 transports	6,10%
340 postes et télécommunications	2,33%
350 services financiers	2,07%
360 activités immobilières	3,32%
370 activités des services aux entreprises	2,74%
380 activités d'administration publique	5,99%
390 éducation et formation	3,64%
400 activités de santé et action sociale	1,23%
410 activités à caractère collectif ou personnel	1,73%
<b>Total</b>	<b>100%</b>

La structure médiane est calculée sur la période de 1980 à 2004 sur les TES en valeur.