



Munich Personal RePEc Archive

The Essentials of VAR Modeling with EViews

Josué, ANDRIANADY and Avo Liantsoa,
RANDRIAMANANTENA

2024

Online at <https://mpra.ub.uni-muenchen.de/122256/>
MPRA Paper No. 122256, posted 02 Oct 2024 06:54 UTC

L'Essentiel des Modèles VAR sous EViews

Josué Ravahiny ANDRIANADY *

Avo Liantsoa RANDRIAMANANTENA

October 1, 2024

Abstract

This manual provides a comprehensive overview of Vector Autoregression (VAR) models using EViews. It covers key concepts, practical applications, and step-by-step guidance on implementing VAR modeling. The aim is to equip researchers and practitioners with the tools necessary to analyze time series data effectively and understand the dynamic relationships among economic variables. By the end of this manual, readers will be able to confidently employ VAR models for their analytical needs.

*Corresponding author: jrahahiny@gmail.com

Avant-Propos

Dans un monde économique en perpétuelle mutation, la maîtrise des outils d'analyse et de prévision des dynamiques temporelles devient une compétence essentielle. C'est avec cet objectif en tête que nous vous présentons ce manuel, dédié à la modélisation autorégressive vectorielle (VAR). Ce guide, à la fois accessible et rigoureux, s'adresse aux chercheurs, étudiants et praticiens désireux de plonger dans les profondeurs des interrelations complexes entre variables économiques.

À travers les pages de ce manuel, vous découvrirez les fondements théoriques et les applications concrètes des modèles VAR, qui se révèlent être des instruments puissants pour appréhender les relations dynamiques entre différentes séries chronologiques. Que vous soyez novice ou déjà initié aux modèles de séries temporelles, nous avons veillé à ce que chaque concept soit présenté de manière claire et didactique, vous permettant ainsi d'acquérir les compétences nécessaires pour naviguer avec aisance dans cet univers.

Nous vous invitons également à consulter l'édition précédente de notre série, qui était consacrée à la régression par moindres carrés ordinaires (OLS). Ce manuel a posé les bases des techniques d'estimation économétriques et représente un passage naturel vers l'exploration des modèles VAR.

<https://mpra.ub.uni-muenchen.de/id/eprint/122199>

En vous engageant dans cette lecture, nous espérons vous offrir des clés pour non seulement enrichir votre compréhension des dynamiques économiques, mais également pour vous encourager à embrasser les défis et opportunités qu'offre l'analyse économétrique.

Nous vous souhaitons une excellente lecture et une exploration enrichissante des interactions économiques à travers le prisme du modèle VAR !

Contents

1	Introduction	4
2	Revue de la littérature	4
3	Estimation sous EViews	5
3.1	Importation des données	5
3.2	Tests de stationnarité	6
3.2.1	Notion de base	6
3.2.2	Application Eviews	8
3.3	Sélection du retard	9
3.4	Tests de Cointégration	10
3.4.1	Notion de base	10
3.4.2	Application Eviews	11
3.5	Modélisation VAR	13
3.5.1	Explication des estimations du modèle de manière simplifiée	14
3.5.2	Choix du retard optimal	15
3.6	Tests de diagnostic	16
3.6.1	Test de la stabilité des racines	16
3.6.2	Cusum test	17
3.7	Residual test	18
3.7.1	Test d'hétéroscédasticité	18
3.7.2	Test d'autocorrélation des résidus	20
3.7.3	Test de normalité	21
3.8	Test de causalité de Granger	22
3.9	Analyse de la Fonction de Réponse Impulsionnelle (IRF)	23

1 Introduction

Les modèles autorégressifs vectoriels (VAR) constituent l'un des outils les plus fondamentaux et influents en économétrie moderne, ayant été introduits par Christopher Sims dans les années 1980 pour surmonter certaines des limitations des modèles économétriques précédents. Dans son travail novateur, Sims (1980) a démontré que ces modèles permettent d'explorer les relations dynamiques entre plusieurs variables économiques, offrant ainsi une vision plus holistique des systèmes économiques complexes. Contrairement aux approches traditionnelles qui traitaient souvent une variable comme exogène, le modèle VAR considère toutes les variables comme endogènes, ce qui reflète plus fidèlement la réalité économique où les variables s'influencent mutuellement au fil du temps (Sims, 1980).

Les applications des modèles VAR sont vastes et variées, allant des prévisions économiques à l'analyse de l'efficacité des politiques monétaires et budgétaires (Luetkepohl, 2011). En intégrant des informations sur les interactions passées entre les variables, ces modèles permettent aux chercheurs et aux décideurs de mieux comprendre les mécanismes sous-jacents aux fluctuations économiques. Par exemple, (Davis et al., 2016) ont souligné la capacité des modèles VAR à identifier les relations causales entre les variables, un aspect crucial pour l'analyse des impacts des chocs économiques sur les économies nationales.

Cependant, malgré leur puissance analytique, les modèles VAR ne sont pas exempts de limitations. La complexité croissante des modèles avec l'augmentation du nombre de variables et de retards peut entraîner des défis en matière d'estimation et d'interprétation (Stock and Watson, 2005).

Ce manuel a pour objectif de fournir une compréhension approfondie des modèles VAR, en abordant leur formulation mathématique, leur estimation, ainsi que leurs applications pratiques. En mettant l'accent sur les meilleures pratiques et les défis potentiels, ce guide aspire à équiper les chercheurs et les praticiens avec les connaissances nécessaires pour utiliser efficacement les modèles VAR dans leurs travaux.

2 Revue de la littérature

Les modèles autorégressifs vectoriels (VAR) sont des modèles économétriques introduits par Christopher Sims dans les années 1980 (Sims, 1980). Ces méthodes statistiques sont utilisées en économétrie pour étudier les relations dynamiques entre plusieurs variables économiques au fil du temps.

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{A}_2 \mathbf{Y}_{t-2} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \mathbf{u}_t$$

où :

- \mathbf{Y}_t est un vecteur de variables endogènes à un temps t .
- \mathbf{c} est un vecteur de constantes (terme d'ordonnée à l'origine).
- \mathbf{A}_i sont des matrices de coefficients pour chaque lag i ($i = 1, 2, \dots, p$).
- \mathbf{u}_t est un vecteur de termes d'erreur (ou chocs) à un temps t , supposés être des variables aléatoires blanches.

Les modèles VAR sont appliqués dans divers domaines. Ils sont notamment utilisés pour effectuer des prévisions économiques (Kina, 1986), pour analyser l'efficacité des politiques monétaires et budgétaires, ainsi que pour comprendre les interactions entre les agents économiques (Luetkepohl, 2011). En outre, ils sont couramment employés dans le secteur financier pour l'étude et la gestion des risques (Pribadi, 2023).

Les modèles VAR présentent plusieurs avantages. Premièrement, ils sont particulièrement efficaces pour fournir des estimations précises dans le cadre de l'analyse économique à court terme (Stock and Watson, 2005). En traitant toutes les variables comme endogènes, ces modèles expliquent les valeurs actuelles des variables par leurs valeurs passées (Kilian, 2013). De plus, ils permettent d'identifier clairement les relations causales entre les variables, ce qui est essentiel pour l'analyse des politiques économiques et la compréhension des chocs économiques (Davis et al., 2016). Par ailleurs, les modèles VAR sont capables de modéliser des interactions complexes entre différentes séries chronologiques, offrant ainsi des prévisions plus précises que celles des modèles ARIMA ou à fonction de transfert. Cela en fait une option solide pour les prévisions économiques (Gao, 2024).

Cependant, les modèles VAR présentent certaines limites. Tout d'abord, leur complexité augmente rapidement avec le nombre de données, ce qui peut entraîner des prévisions et des estimations instables en raison du nombre élevé de coefficients autoregressifs (Davis et al., 2016). De plus, l'estimation des facteurs dynamiques peut être problématique, car les études empiriques privilégient souvent des modèles approximatifs plutôt que des modèles exacts (Stock and Watson, 2005), ce qui rend l'identification des chocs structurels plus difficile. Enfin, l'hypothèse selon laquelle toutes les variables sont endogènes ne reflète pas toujours la réalité des relations économiques. Dans ce cas, les modèles VAR peuvent fournir des interprétations erronées si les hypothèses d'exogénéité ne sont pas bien fondées (Gao, 2024).

Bien que les modèles VAR soient largement utilisés, il existe des alternatives pour modéliser les relations entre variables économiques. Par exemple, les modèles à correction d'erreur vectorielle (VECM) sont souvent utilisés lorsque les variables étudiées sont co-intégrées, c'est-à-dire lorsqu'elles partagent une relation d'équilibre à long terme. Le VECM permet de capturer les dynamiques à court terme tout en prenant en compte les ajustements vers cet équilibre à long terme (Lütkepohl, 2004).

Une autre alternative est l'utilisation des modèles à facteurs dynamiques (DFM), qui réduisent la dimensionnalité des données en résumant un grand nombre de séries temporelles en un nombre restreint de facteurs communs (Stock and Watson, 2011). Ces modèles sont particulièrement efficaces pour l'analyse de données complexes ou avec un grand nombre de variables, tout en réduisant la sur-paramétrisation observée dans les modèles VAR classiques.

Enfin, les modèles de régression pénalisée, tels que le LASSO-VAR (Belloni et al., 2010), sont une autre approche moderne qui impose des contraintes de parcimonie sur les coefficients pour éviter le sur-ajustement des modèles VAR lorsqu'un grand nombre de variables est impliqué.

3 Estimation sous EViews

3.1 Importation des données

- **Préparation des données** : Avant de commencer, assurez-vous que vos données sont dans un format compatible avec EViews. Les formats les plus couramment utilisés incluent CSV (Comma-Separated Values) et Excel (XLS ou XLSX). Les séries doivent être bien structurées, avec les dates en colonnes et les valeurs en lignes. Voici quelques conseils pour la préparation des données :
 - Chaque série temporelle doit avoir une colonne dédiée pour les dates. Les dates doivent être formatées de manière cohérente (par exemple, YYYY-MM-DD).
 - Les valeurs correspondantes doivent être placées dans les colonnes adjacentes, sans espaces vides entre les données.
 - Si vous utilisez un fichier CSV, assurez-vous qu'il n'y a pas de caractères spéciaux ou d'en-têtes non nécessaires dans le fichier qui pourraient perturber l'importation.

- **Utiliser la commande `wfopen`** : Une fois vos données préparées, vous pouvez les ouvrir dans EViews en utilisant la commande suivante :

```
wfopen "chemin_du_fichier"
```

- La commande `wfopen` est utilisée pour ouvrir un fichier de travail dans EViews. Cette commande charge les données du fichier spécifié dans un nouvel espace de travail.
- Remplacez "`chemin_du_fichier`" par le chemin d'accès réel à votre fichier. Par exemple, si votre fichier s'appelle `data.csv` et se trouve sur votre bureau, cela pourrait ressembler à :

```
wfopen "C:\Users\Desktop\data.csv"
```

- Assurez-vous d'utiliser des guillemets autour du chemin du fichier pour éviter les erreurs d'interprétation.

- **Vérification des données** : Après avoir importé les données, il est essentiel de vérifier que toutes les séries sont correctement chargées. Pour ce faire, utilisez la commande suivante pour afficher le contenu du fichier de travail :

```
show
```

- La commande `show` ouvre une fenêtre d'affichage qui présente toutes les séries de données importées.
- Vérifiez que les séries affichées correspondent à celles présentes dans votre fichier source. Cela vous permettra de confirmer que les données ont été importées sans erreur et que leur structure est conforme à vos attentes.
- Si des anomalies sont détectées (comme des valeurs manquantes ou des données mal alignées), il peut être nécessaire de retourner à votre fichier source pour corriger les problèmes avant de réimporter les données.

3.2 Tests de stationnarité

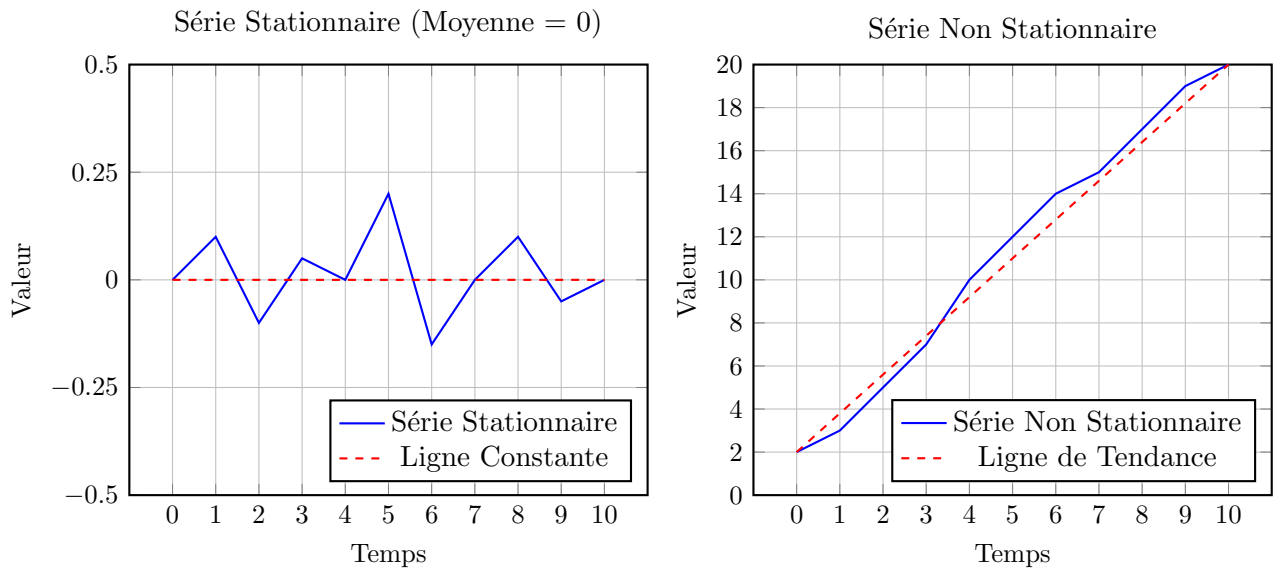
3.2.1 Notion de base

Pourquoi effectuer un test de stationnarité?

Faire un test de stationnarité est essentiel pour s'assurer que les données restent constantes au fil du temps. Si une série temporelle est stationnaire (la série est toujours constante au fil du temps), cela signifie que ses caractéristiques, comme la moyenne et la variance, ne changent pas, ce qui permet d'utiliser les comportements passés pour faire des prédictions futures. Les modèles de type VAR, qui analysent ces données, fonctionnent mieux avec des séries stationnaires. Si les données ne le sont pas, cela peut fausser les résultats et rendre la modélisation et les prévisions moins fiables. En vérifiant la stationnarité, on s'assure donc que les analyses sont solides et que les conclusions sont valides.

Les deux graphiques illustrent la distinction entre une série temporelle stationnaire et une série temporelle non stationnaire. Dans le premier graphique, la série stationnaire oscille autour d'une valeur moyenne, représentée par une ligne rouge, avec des fluctuations régulières qui indiquent que ses propriétés statistiques, telles que la moyenne et la variance, restent constantes dans le temps. Cela rend cette série utile pour les analyses statistiques, car elle permet des prévisions plus fiables.

En revanche, le second graphique montre une série non stationnaire avec une tendance croissante continue, illustrée par une ligne rouge qui s'élève avec le temps. Les valeurs augmentent de manière systématique, ce qui signifie que les caractéristiques statistiques de cette série varient au fil du temps.



(a) Série Stationnaire avec Moyenne de Zéro

(b) Série Non Stationnaire

Notion de racine unitaire

La racine unitaire est une caractéristique qui tire une série vers une tendance non stationnaire. Le graphique ci dessous compare deux types de séries temporelles. La première, représentée par une ligne bleue, illustre une série avec racine unitaire, qui montre une tendance à la hausse continue. Cela signifie que les valeurs de cette série augmentent sans jamais se stabiliser, et les flèches bleues vers le haut indiquent comment les racines unitaires attirent les valeurs vers des niveaux de plus en plus élevés.

En revanche, la deuxième série, illustrée par une ligne rouge, est stationnaire, oscillant autour d'une valeur moyenne constante, visible par la ligne pointillée rouge. Les flèches rouges montrent que, bien que cette série fluctue, elle reste proche de sa moyenne, ce qui indique une certaine stabilité.

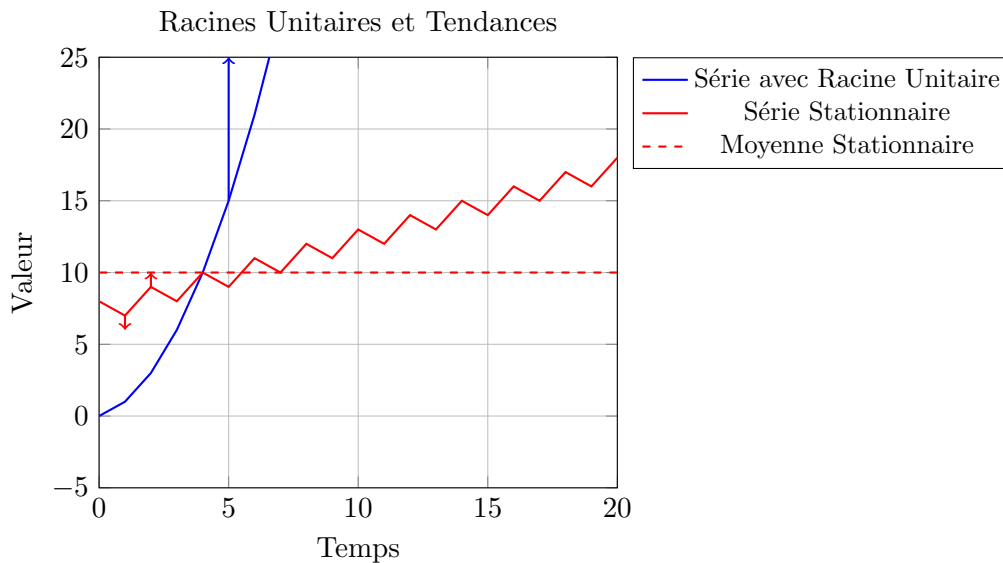


Figure 1: Illustration des racines unitaires montrant une série non stationnaire (en bleu) et une série stationnaire (en rouge) .

Il existe plusieurs tests de racine unitaire couramment utilisés pour déterminer la stationnarité des séries temporelles.

Le test d'Augmented Dickey-Fuller (ADF) est le plus répandu ; il teste l'hypothèse nulle selon laquelle une série est non stationnaire, en intégrant des retards pour prendre en compte l'autocorrélation dans les résidus.

Le test de Phillips-Perron (PP) est une variante qui corrige les problèmes d'hétéroscédasticité et d'autocorrélation, sans inclure de retards supplémentaires.

En revanche, le test de Kwiatkowski-Phillips-Schmidt-Shin (KPSS) teste l'hypothèse nulle de stationnarité, offrant ainsi une vérification alternative aux résultats des tests ADF et PP.

Pour des séries présentant des ruptures structurelles, le test de Zivot-Andrews est particulièrement utile, tandis que le test de Ng-Perron améliore la puissance de détection des racines unitaires dans des séries ayant des tendances ou des ruptures. Le choix du test dépend des caractéristiques spécifiques de la série temporelle analysée.

3.2.2 Application Eviews

- **Effectuer le test ADF** : Pour chaque série, utilisez la méthode Augmented Dickey-Fuller (ADF) afin d'évaluer la stationnarité, deux alternatives de commandes :

```
y1.uroot
nom_variable.uroot
```

- `y1` représente vos séries de données que vous souhaitez analyser.
 - La commande `.uroot` effectue le test ADF avec 0 retards et une tendance constante. Cette spécification est appropriée lorsque vous supposez que la série peut contenir une tendance mais sans composante saisonnière.
 - Les résultats du test vous indiqueront si vous pouvez rejeter l'hypothèse nulle, qui stipule que la série est non stationnaire. Si la valeur p associée au test est inférieure à un seuil significatif (habituellement 0,05), cela suggère que vous pouvez rejeter l'hypothèse nulle.
- **Interpréter les résultats du test** : Après avoir exécuté le test ADF, examinez les p-values des tests pour chacune des séries :
 - Si la p-value est inférieure à 0,05, vous pouvez rejeter l'hypothèse nulle. Cela signifie que la série est considérée comme stationnaire, et vous pouvez poursuivre votre analyse avec cette série.
 - En revanche, si la p-value est supérieure à 0,05, vous ne pouvez pas rejeter l'hypothèse nulle, ce qui indique que la série est non stationnaire. Dans ce cas, vous devrez appliquer une transformation pour rendre la série stationnaire.

Null Hypothesis: Y1 has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on SIC, maxlag=7)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.564165	0.7832
Test critical values:		
1% level	-4.296729	
5% level	-3.568379	
10% level	-3.218382	

*MacKinnon (1996) one-sided p-values.

Figure 2: Illustration test de stationnarité ADF (variable non stationnaire)

- **Différencier les séries non stationnaires** : Si une ou plusieurs séries ne sont pas stationnaires, vous devez créer leurs différences pour tenter de les rendre stationnaires :

```
series d_y1 = d(y1)
gen d_nom_variable = d(y2)
```

- La commande `series` est utilisée pour créer une nouvelle série à partir des séries existantes. Ici, `d(y1)` représente la première différence de `y1`.
- La première différence d'une série temporelle est calculée en soustrayant la valeur précédente de la valeur actuelle. Cela peut aider à éliminer les tendances et rendre les séries stationnaires.
- Une fois que vous avez créé les séries différenciées `d_y1` vous pouvez répéter le test ADF pour vérifier à nouveau la stationnarité.

3.3 Sélection du retard

Dans un modèle VAR (Vector Autoregression), le retard fait référence au nombre de périodes passées que l'on inclut pour prédire les valeurs futures d'une série temporelle. Le retard par défaut est 2, ensuite on estime le modèle VAR puis on refais le test de longueur pour valider le retard optimal. Par exemple, si nous choisissons un retard de 2, cela signifie que nous utilisons les valeurs des deux périodes précédentes pour prévoir la prochaine valeur.

Pourquoi choisir un retard de 2 par défaut ?

- **Équilibre entre Simplicité et Précision** : un retard de 2 permet de capturer les effets immédiats et ceux qui se manifestent après un certain temps. Cela aide à mieux comprendre comment une variable influence une autre.
- **Norme Établie** : de nombreuses études économiques utilisent un retard de 2, ce qui facilite la comparaison de nos résultats avec ceux d'autres recherches.
- **Tests Statistiques Favorables** : les critères d'information, qui aident à évaluer la qualité du modèle, montrent souvent qu'un retard de 2 est optimal pour de nombreuses séries de données.
- **Effets Différés** : dans le monde réel, les changements dans une variable (comme le prix du pétrole) peuvent ne pas avoir d'impact immédiat sur l'économie. Un retard de 2 aide à prendre en compte ces effets retardés.

3.4 Tests de Cointégration

3.4.1 Notion de base

Qu'est-ce que la Cointégration ?

La cointégration est une méthode statistique utilisée pour déterminer si deux ou plusieurs séries temporelles (comme par exemple le PIB et le taux d'inflation) ont une relation stable à long terme. Même si ces séries montrent des variations ou des tendances (c'est-à-dire qu'elles ne sont pas stationnaires), elles peuvent se déplacer ensemble de manière cohérente.

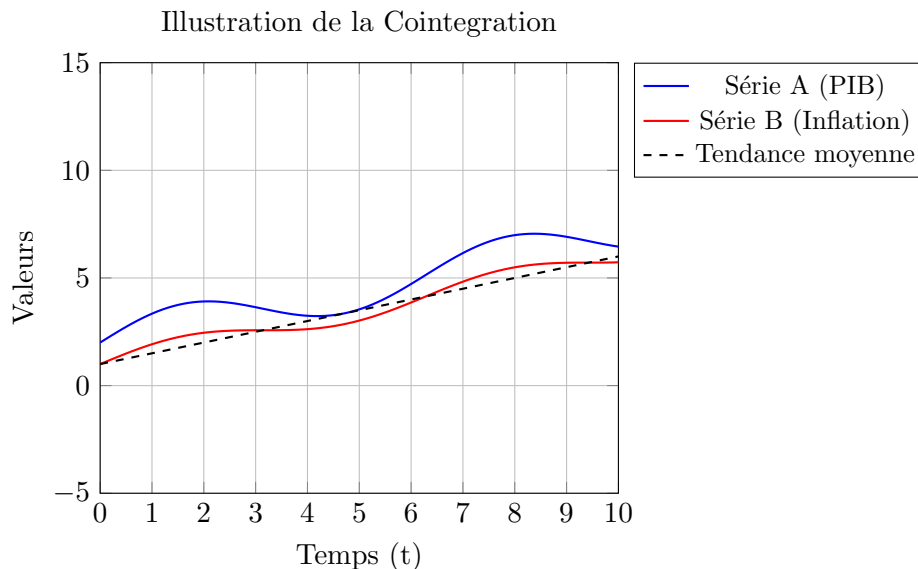


Figure 3: Illustration en graphe de la cointégration.

Bien que le PIB et le taux d'inflation puissent fluctuer de manière irrégulière au fil du temps, la cointégration indique qu'il existe une relation stable. Par exemple, si le PIB augmente, il peut y avoir une tendance à ce que le taux d'inflation augmente aussi, même si cela ne se produit pas immédiatement.

Pourquoi c'est important ?

- **Relations Stables** : La cointégration permet de comprendre comment des variables économiques interagissent sur le long terme. Par exemple, si le PIB et le taux d'inflation sont cointégrés, cela suggère qu'ils évoluent de manière liée, ce qui peut être essentiel pour les décisions de politique économique.
- **Modèles Avancés** : Lorsque des séries sont cointégrées, il est possible d'utiliser des modèles plus sophistiqués, tels que les modèles de correction d'erreurs (VECM). Ces modèles sont plus précis car ils tiennent compte de la dynamique à court terme et de la relation à long terme entre les variables.
- **Test de Johansen** : Ce test est essentiel pour déterminer s'il existe des relations de cointégration entre plusieurs séries temporelles. Une relation de cointégration indique qu'il existe une relation stable à long terme entre les variables, même si les séries elles-mêmes ne sont pas stationnaires.

3.4.2 Application Eviews

Commandes pour le Test de Cointegration

Entrez la commande suivante dans la fenêtre de commande :

```
coint(johansen, 2) d_y1 d_x
```

- `coint` : Indique que vous réalisez un test de cointegration.
- `johansen` : Spécifie que le test est de type Johansen.
- `2` : Représente le nombre de retards à inclure (choisi en fonction des résultats de vos tests de stationnarité).
- `d_y1` et `d_x` : Les deux séries que vous testez pour la cointegration.

Interprétation des Résultats

Après avoir exécuté la commande, EViews affichera les résultats. Voici les éléments clés à surveiller :

Valeurs de Trace :

Le tableau affichera une série de valeurs de trace pour chaque rang. Vous devez les comparer avec les valeurs critiques de la statistique de trace pour différents niveaux de signification (généralement 1%, 5% et 10%).

Rang	Valeur de Trace	Valeur Critique (5%)	Statistique de Probabilité
0	32.50	20.25	0.0001
1	15.75	9.14	0.0015

Table 1: Résultats des Valeurs de Trace

Si la valeur de trace est supérieure à la valeur critique, cela suggère qu'il existe au moins une relation de cointegration. Dans cet exemple, la valeur de trace pour le rang 0 (32.50) est supérieure à la valeur critique (20.25), indiquant une relation de cointegration significative.

Valeurs de Max-Eigenvalue

Ces valeurs indiquent également le nombre de relations de cointegration et doivent être comparées aux valeurs critiques.

Rang	Valeur de Max-Eigenvalue	Valeur Critique (5%)	Statistique de Probabilité
0	16.75	15.67	0.0240

Table 2: Résultats des Valeurs de Max-Eigenvalue

Comme pour la valeur de trace, si la valeur de Max-Eigenvalue est supérieure à la valeur critique, cela indique une relation de cointegration. Ici, pour le rang 0, la valeur de Max-Eigenvalue (16.75) est supérieure à la valeur critique (15.67), ce qui renforce la preuve de cointegration.

Examinez également les p-values associées aux tests. Une p-value inférieure à 0.05 indique que la relation de cointégration est significative au niveau de confiance de 95%.

Conclusion du choix du modèle VECM ou VAR

- Si le test indique qu'il existe des relations de cointégration, cela suggère qu'un modèle de correction d'erreurs vectoriel (VECM) est approprié pour modéliser les dynamiques à court terme tout en tenant compte des relations à long terme entre les variables.
- Un VECM peut vous aider à capturer les effets de court terme et à ajuster pour toute déviation à long terme par rapport à l'équilibre.
- En revanche, sans relations de cointégrations, on va toujours opter pour le modèle VAR.

Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.472046	25.97191	15.49471	0.0009
At most 1 *	0.276153	8.725744	3.841465	0.0031

Trace test indicates 2 cointegrating eqn(s) at the 0.05 level
 * denotes rejection of the hypothesis at the 0.05 level
 **MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)				
Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.472046	17.24617	14.26460	0.0164
At most 1 *	0.276153	8.725744	3.841465	0.0031

Max-eigenvalue test indicates 2 cointegrating eqn(s) at the 0.05 level
 * denotes rejection of the hypothesis at the 0.05 level
 **MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b*S11*b=I):

DY	DX
-0.165559	-0.004328
0.135120	-0.009350

Figure 4: Illustration du test sous Eviews

3.5 Modélisation VAR

```
var var_model.ls
```

- *var* indique que vous créez un modèle VAR.
- *var_model* est le nom que vous donnez au modèle.

The figure shows two side-by-side screenshots of the EViews software interface for model specification. Both screenshots show the 'VAR type' dropdown menu, the 'Estimation sample' field, the 'Lag Intervals for Endogenous' field, and the 'Exogenous variables' field. In screenshot (a), the 'VAR type' is set to 'Standard VAR'. In screenshot (b), the 'VAR type' is set to 'Vector Error Correction'. The 'Estimation sample' is '1990 2020', the 'Lag Intervals for Endogenous' is '1 2', and the 'Exogenous variables' is 'c' in both.

(a) Choisir cette option si modèle VAR

(b) Choisir cette option si modèle VECM

Figure 5: Choix entre l'estimation VAR et VECM

- **Examiner les résultats de l'estimation** : Après l'estimation, EViews affichera les coefficients, les erreurs standard et les statistiques t. Vérifiez la signification des coefficients. Cela implique :
 - **t-statistiques** : Vérifiez si les valeurs absolues des t-statistiques sont supérieures à 2 (ou en fonction du niveau de signification choisi, généralement 1.96 pour un test à 95% de confiance). Ils sont généralement dans un crochet []. Cela indique si les coefficients sont statistiquement significatifs.
 - **Interprétation des Coefficients** : Regardez les coefficients eux-mêmes pour comprendre l'effet des variables sur la variable dépendante. Par exemple, un coefficient négatif signifie qu'une augmentation de la variable indépendante est associée à une diminution de la variable dépendante.
 - **R-squared** : Indique la proportion de variance de la variable dépendante expliquée par le modèle. Un R-squared faible suggère que le modèle n'explique pas bien les données.
 - **Adjusted R-squared** : Plus pertinent que R-squared, car il prend en compte le nombre de variables dans le modèle. Un ajustement négatif peut indiquer un problème avec le modèle.
 - **F-statistique** : Vérifiez si la statistique F est significative. Cela indique si au moins un des coefficients dans le modèle est significativement différent de zéro.

Table 3: Illustration résultat VAR sous Eviews

Variables	DY	DX
DY(-1)	-0.234567 (0.12500) [-1.87654]	-2.456789 (0.51234) [-4.80347]
DY(-2)	-0.156789 (0.09876) [-1.58765]	0.987654 (0.34567) [2.85714]
DX(-1)	-0.012345 (0.00567) [-2.18344]	0.022345 (0.01456) [1.53274]
DX(-2)	-0.004567 (0.00123) [-3.70378]	-0.045678 (0.01234) [-2.87654]
Constante (C)	1.123456 (0.51234) [2.19456]	150.9876 (62.1234) [2.42689]
R-squared	0.750198	0.265657
Adj. R-squared	0.740345	0.180345
Sum sq. resids	1234.567	987654.3
S.E. equation	9.87654	215.4321
F-statistic	5.678912	4.789012
Log likelihood	-120.234	-190.567
Akaike AIC	6.789012	14.34567
Schwarz SC	7.012345	14.56789
Mean dependent	-0.456789	130.4321
S.D. dependent	10.12345	198.7654
Determinant resid covariance (dof adj.)	3500000.	
Determinant resid covariance	2500000.	
Log likelihood	-295.6789	
Akaike information criterion	19.01234	
Schwarz criterion	19.56789	
Nombre de coefficients	10	

3.5.1 Explication des estimations du modèle de manière simplifiée

- Coefficient (en écroché) :

- Ce chiffre représente l'impact d'une variable sur une autre.
- *Exemple* : Pour le coefficient de $DY(-1)$ affiché comme **-0.023**, cela signifie qu'une augmentation d'une unité de $DY(-1)$ entraîne une diminution de 0.023 unités de DY .

- Erreurs standards (en parenthèses) :

- Ce chiffre, qui est affiché en parenthèses, mesure la précision du coefficient estimé. Une erreur standard plus petite indique une estimation plus précise.
- *Exemple* : Pour $DY(-1)$, l'erreur standard est (**0.125**). Cela signifie que l'incertitude autour de l'estimation du coefficient est de 0.125.

- **t-statistiques (en crochets) :**

- Ce chiffre, qui apparaît entre crochets, est utilisé pour tester la significativité du coefficient. Un t-statistique supérieur à 2 (ou inférieur à -2) indique que le coefficient est significatif.
- *Exemple* : Pour $DY(-1)$, la t-statistique est [-1.87]. Dans ce cas, elle n'est pas significative, mais un t-statistique supérieur à 2 signifierait que le coefficient est significatif.

3.5.2 Choix du retard optimal

Dans EViews, utilisez la commande suivante pour accéder aux critères de sélection de la longueur des retards :

```
var_model.laglen(2)
```

- *var_model* est le nom que vous donnez au modèle.
- *laglen* fait appel au commande lag length criteria pour laisser Eviews choisir le retard optimal.
- (2) retard par défaut défini précédemment.

Critères de Sélection de la Longueur des Retards avec Étoiles sous EViews

Lorsqu'on sélectionne la longueur des retards d'un modèle VAR sous EViews, les étoiles (*) indiquent le nombre optimal de retards suggéré par chaque critère. Voici un exemple de sortie des résultats pour la sélection des retards.

Exemple de Résultats

```
VAR Lag Order Selection Criteria
Endogenous variables: Y X
Exogenous variables: C
Sample: 1 100
Included observations: 96
```

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-10.32	NA	0.002	2.63	2.82	2.70
1	25.45	15.23	0.0005	1.85*	2.16	1.95
2	45.76	9.14	0.0003	1.56	2.10	1.80
3	60.89	6.43	0.0002*	1.28	2.05*	1.67*

Interprétation des Résultats

Les étoiles (*) signalent les critères qui minimisent les valeurs et indiquent la meilleure longueur de retard :

- **AIC (Akaike Information Criterion)** : La plus petite valeur de 1.28 se trouve au retard 3, ce qui suggère que le modèle avec 3 retards est optimal selon l'AIC.
- **SC (Schwarz Criterion)** : Le critère SC choisit 1 retard comme le meilleur modèle, car c'est là que la valeur est la plus basse (2.16).

- **HQ (Hannan-Quinn Criterion)** : Le HQ recommande aussi 3 retards (valeur de 1.67), marqué avec une étoile.
- **FPE (Final Prediction Error)** : Le FPE est également le plus faible au retard 3, avec une valeur de 0.0002, ce qui est également marqué par une étoile.

Ainsi, pour ce modèle VAR, il semble que les critères AIC, HQ et FPE suggèrent tous d'utiliser 3 retards, tandis que SC recommande 1 retard.

Estimation du VAR avec la longueur de retard sélectionnée

Une fois le nombre optimal de retards déterminé (ici, 3 retards), utilisez la commande suivante pour estimer le modèle VAR avec 3 retards :

```
var var_model.ls 1 3 Y X
```

Cette commande estime le modèle VAR pour Y et X avec 3 retards.

3.6 Tests de diagnostic

Ici on va vérifier la validité du modèle en effectuant plusieurs tests.

3.6.1 Test de la stabilité des racines

Dans un modèle VAR, un test de stabilité est réalisé en vérifiant si les racines caractéristiques du modèle sont à l'intérieur du cercle unité. Si c'est le cas, le modèle est considéré comme stable. Pour réaliser ce test dans EViews, utilisez la commande suivante après avoir estimé votre VAR :

```
Nom_model.arroots(graph)
```

- *Nom_model* est le nom que vous donnez au modèle.
- **arroots** fait appel à la commande pour faire le test stabilité des racines
- (**graph**) donne le cercle de test de stabilité.

Cette commande vous permet d'obtenir les racines caractéristiques et de visualiser leur position par rapport au cercle unité. Le test confirme si le modèle VAR est stable ou non.

Voici une illustration de ce test sous forme de graphe.

Dans le graphique ci-dessus :

- Le cercle bleu représente le cercle unité.
- Les points rouges représentent les racines caractéristiques de votre modèle VAR.
- Si toutes les racines se trouvent à l'intérieur du cercle (comme ici), cela signifie que le modèle VAR est stable.

Cercle Unitaire et Racines Caractéristiques

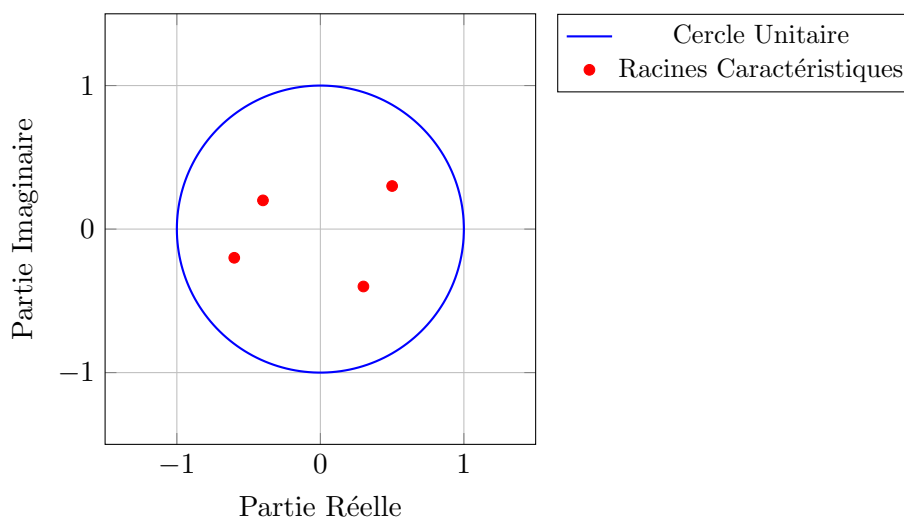


Figure 6: Les racines caractéristiques à l'intérieur du cercle unité indiquent que le modèle VAR est stable.

3.6.2 Cusum test

Qu'est-ce que le Test CUSUM ?

Le test CUSUM (Cumulative Sum Control Chart) est utilisé pour vérifier si les coefficients d'un modèle de régression restent constants au fil du temps. Cela permet d'identifier si le modèle est toujours valable ou s'il y a eu des changements dans la relation entre les variables.

Commande:

```
rls
```

Puis choisir CUSUM TEST et CUSUM SQUARED

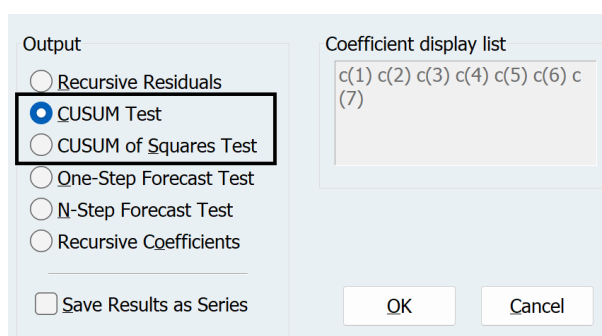


Figure 7: Choix du test cusum sous Eviews

Si la courbe CUSUM reste entre les lignes de limite(en rouge), cela signifie que le modèle est stable. Les relations entre les variables n'ont pas changé. Au-dessus de la Limite Supérieure :

Si la courbe dépasse la limite, cela indique que le modèle pourrait être instable. Il se peut qu'il ne fonctionne plus bien.

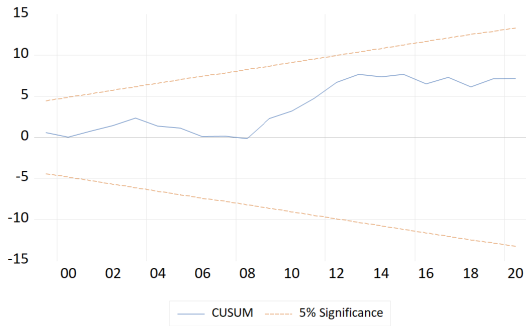


Figure 8: Test CUSUM

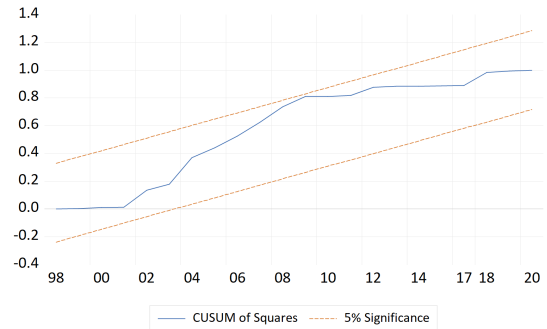


Figure 9: Test CUSUM Squared

Figure 10: Résultats des Tests CUSUM et CUSUM Squared

3.7 Residual test

3.7.1 Test d'hétéroscédasticité

L'hétéroscédasticité se produit lorsque les erreurs d'un modèle de régression (les écarts entre les valeurs prédites et réelles) varient en fonction des variables utilisées. En d'autres termes, les erreurs peuvent être plus grandes ou plus petites selon les valeurs des variables, ce qui peut rendre les résultats moins précis et moins fiables.

La commande est :

```
Nom_modele.white
```

- L'argument `Nom_modele` fait référence aux nom de l'équation VAR spécifié.
- `.white` commande qui fait appel aux test de White est un test statistique utilisé pour détecter l'hétéroscédasticité dans un modèle de régression.

Joint test:		
Chi-sq	df	Prob.
190.0102	180	0.2901

Individual components:					
Dependent	R-squared	F(12,13)	Prob.	Chi-sq(12)	Prob.
res1*res1	0.431342	0.821736	0.6299	11.21489	0.5106
res2*res2	0.489907	1.040462	0.4696	12.73758	0.3884
res3*res3	0.718103	2.759677	0.0408	18.67068	0.0968
res4*res4	0.361542	0.613464	0.7972	9.400104	0.6684
res5*res5	0.596244	1.599805	0.2061	15.50234	0.2151
res2*res1	0.700128	2.529316	0.0552	18.20332	0.1097
res3*res1	0.297935	0.459733	0.9058	7.746301	0.8046
res3*res2	0.699860	2.526090	0.0554	18.19635	0.1099
res4*res1	0.448037	0.879358	0.5850	11.64896	0.4743
res4*res2	0.497564	1.072828	0.4485	12.93666	0.3737
res4*res3	0.384559	0.676921	0.7467	9.998528	0.6161
res5*res1	0.449831	0.885760	0.5801	11.69562	0.4704
res5*res2	0.655197	2.058557	0.1055	17.03512	0.1483
res5*res3	0.764743	3.521561	0.0162	19.88332	0.0693
res5*res4	0.368496	0.632149	0.7825	9.580903	0.6527

Figure 11: Illustration résultat sous Eviews

- Si la p-value est supérieure à 0,05 :

- Vous n’avez pas suffisamment de preuves pour rejeter l’hypothèse nulle.
- *Conclusion* : Il n’y a pas de problème d’hétéroscédasticité ; la variance des erreurs est constante.

- Si la p-value est inférieure ou égale à 0,05 :

- Vous rejetez l’hypothèse nulle.
- *Conclusion* : Il existe des preuves d’hétéroscédasticité ; la variance des erreurs n’est pas constante. Cela pourrait affecter la fiabilité des résultats du modèle, notamment les estimations des coefficients et les tests de significativité.

3.7.2 Test d'autocorrélation des résidus

Qu'est-ce que c'est ?

Le test d'autocorrélation des résidus est un outil qui permet de vérifier si les erreurs d'un modèle de prévision (les résidus) sont liées entre elles. Autrement dit, il permet de savoir si les erreurs que nous avons faites dans le passé affectent celles que nous faisons maintenant ou dans le futur.

Pourquoi est-ce important ?

- **Indépendance des Résidus** : Pour qu'un modèle soit considéré comme fiable, il est crucial que les résidus soient indépendants. Si les erreurs sont corrélées, cela suggère que le modèle ne parvient pas à saisir pleinement la dynamique des données, indiquant ainsi la nécessité d'une amélioration du modèle.
- **Diagnostic de Modèle** : La détection d'autocorrélations significatives dans les résidus peut signaler que le modèle omet certaines variables importantes ou qu'il nécessite une structure différente pour mieux représenter les données.

Commande dans Eviews

`Nom_de_equation.arlm(2)`

- *Nom_de_equation* cette commande indique le nom de l'équation du modèle VAR à estimer
- *.arlm* commande qui fait appel aux tests de corrélation
- *(2)* nombre de retard

Null hypothesis: No serial correlation at lag h						
Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	15.71331	25	0.9231	0.568879	(25, 38.7)	0.9301
2	21.24054	25	0.6791	0.815298	(25, 38.7)	0.7011

Null hypothesis: No serial correlation at lags 1 to h						
Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	15.71331	25	0.9231	0.568879	(25, 38.7)	0.9301
2	54.79700	50	0.2976	1.036792	(50, 26.2)	0.4724

*Edgeworth expansion corrected likelihood ratio statistic.

Figure 12: Illustration du test sous Eviews

- **Si la p-value (Prob) est faible** (en général, inférieure à un seuil de 0,05), vous rejetez l'hypothèse nulle et concluez qu'il y a une autocorrélation significative des résidus. Cela signifie que les erreurs du modèle sont corrélées dans le temps, ce qui pourrait indiquer que le modèle n'a pas capturé toutes les dynamiques des données.
- **Si la p-value (Prob) est élevée** (supérieure à 0,05), vous ne rejetez pas l'hypothèse nulle, ce qui suggère qu'il n'y a pas de corrélation significative entre les résidus. Les erreurs sont donc considérées comme indépendantes, ce qui est un bon signe pour la validité du modèle.

3.7.3 Test de normalité

Le test de normalité des résidus vérifie si les erreurs de prévision dans un modèle de régression suivent une distribution normale. La normalité des résidus est importante car elle garantit la fiabilité des résultats du modèle (coefficients, tests t et F).

Pourquoi tester la normalité des résidus ?

- Si les résidus ne suivent pas une distribution normale, cela peut indiquer que le modèle est mal spécifié ou que certaines variables explicatives importantes ont été omises.
- Des résidus non normaux peuvent entraîner des estimations biaisées des coefficients et des tests de significativité incorrects.

Comment interpréter les résultats ?

- Si la **p-value est supérieure à 0,1** : L'hypothèse de normalité des résidus est acceptée. Les résidus sont normalement distribués.
- Si la **p-value est inférieure à 0,1** : L'hypothèse de normalité est rejetée. Les résidus ne suivent pas une distribution normale, ce qui peut poser un problème pour l'analyse.

Commande dans EViews

var_model.jbera

Component	Skewness	Chi-sq	df	Prob.*
1	0.017108	0.001687	1	0.9672
2	1.085834	6.794568	1	0.0091
Joint		6.796255	2	0.0334

Component	Kurtosis	Chi-sq	df	Prob.
1	3.070592	0.151663	1	0.6970
2	4.064427	3.183475	1	0.0744
Joint		3.335137	2	0.1887

Component	Jarque-Bera	df	Prob.
1	0.153349	2	0.9262
2	9.978043	2	0.0068
Joint	11.63670	9	0.2346

*Approximate p-values do not account for coefficient estimation



Figure 13: Illustration résultat du test de normalité sous Eviews

3.8 Test de causalité de Granger

Le test de causalité de Granger est une méthode utilisée en économétrie pour déterminer si une variable peut prédire une autre variable dans une série temporelle. Cette méthode repose sur l'idée que si une variable X Granger-cause une variable Y , alors les valeurs passées de X contiennent des informations qui aident à prédire les valeurs futures de Y .

Commande dans Eviews

```
group Nom_du_groupe Y1 X
Nom_du_groupe.cause
```

- *groupe* : Cette commande crée un groupe de variables nommé *Nom_du_groupe*, qui contient les variables $Y1$ et X .
- *.cause* : Après avoir défini le groupe, cette commande exécute le test de causalité de Granger pour les variables dans le groupe créé.

Les résultats fictifs du test de causalité de Granger entre les variables Y et X sont présentés dans le tableau ci-dessous :

Table 4: Résultats du Test de Causalité de Granger

Hypothèse Nulle	Obs	F-Statistic	Probabilité
X ne Granger-cause pas Y	26	1.50680	0.2446
Y ne Granger-cause pas X	26	0.30953	0.7371

- La première ligne indique que l'hypothèse selon laquelle X ne Granger-cause pas Y n'est pas rejetée, car la valeur p est supérieure à 0.05. Cela signifie que les valeurs passées de X ne prédisent pas significativement Y .
- La deuxième ligne montre que l'hypothèse selon laquelle Y ne Granger-cause pas X n'est également pas rejetée. La valeur p étant bien supérieure à 0.05 indique qu'il n'y a pas de preuve que les valeurs passées de Y prédisent X .
- Le test de causalité de Granger a révélé qu'il n'y a pas de relation de causalité significative entre les variables Y et X . Cela peut suggérer que les deux séries évoluent de manière indépendante ou que d'autres facteurs influencent ces variables.

3.9 Analyse de la Fonction de Réponse Impulsionnelle (IRF)

La fonction de réponse impulsionnelle (IRF) mesure l'effet d'un choc unitaire sur une variable endogène dans un modèle VAR (Vector AutoRegression). Plus précisément, elle indique la réponse dynamique de chaque variable dans le système à un choc sur une autre variable, tout en maintenant constantes toutes les autres innovations.

Les graphiques de la réponse impulsionnelle montrent la réponse d'une variable à un choc sur une autre variable à travers différentes périodes :

- **Si la réponse est positive**, cela signifie que la variable réagit en augmentant à la suite du choc.
- **Si la réponse est négative**, cela signifie que la variable diminue en réponse au choc.
- **La durée de l'effet est également importante** : une réponse impulsionnelle persistante peut indiquer une dynamique à long terme, tandis qu'une réponse qui se dissipe rapidement indique un effet de court terme.

La commande suivante génère une fonction de réponse impulsionnelle dans un modèle VAR

```
Nom_modele.impulse
```

- L'argument `Nom_modele` fait référence aux nom de l'équation VAR spécifié.
- `.impulse` Cette commande montre la réponse de la variable X à un choc sur la variable Y pour une période de 10 unités de temps.

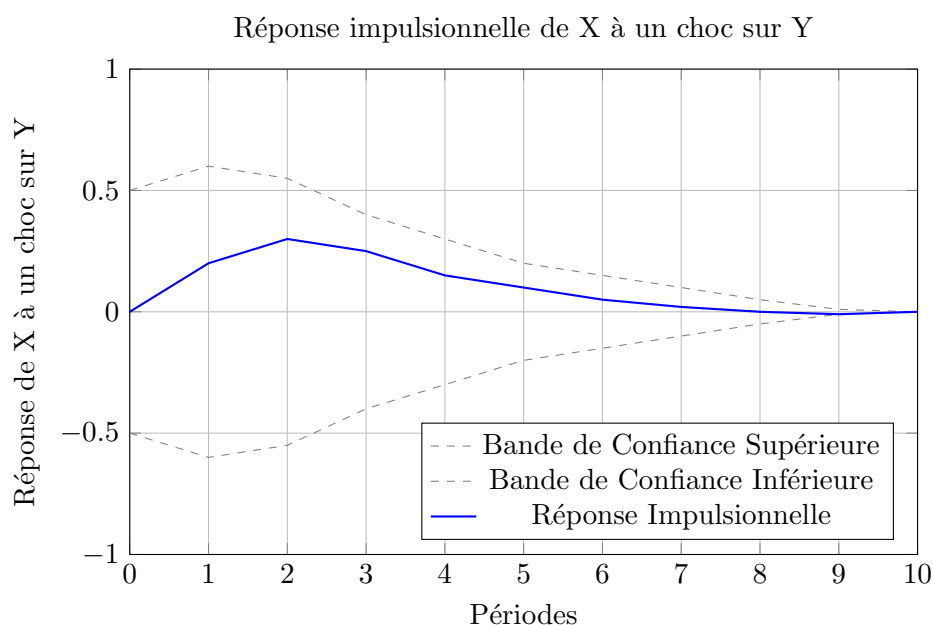


Figure 14: Illustration Réponse impulsionnelle de X à un choc sur Y sous Eviews

References

- Belloni, A., Chernozhukov, V., & Hansen, C. (2010). LASSO methods for Gaussian instrumental variables models. arXiv: Methodology.
- Christopher Sims, "Macroeconomics and Reality", *Econometrica*, Vol. 48, No. 1, 1980, pp. 1-48.
- Davis, R. A., Zang, P., & Zheng, T. (2016). Sparse Vector Autoregressive Modeling. *Journal of Computational and Graphical Statistics*, 25(4), 1077–1096. <https://doi.org/10.1080/10618600.2015.1092978>
- F. Pribadi, A. Surwanti, and W.-C. Shih, "VaR Model for Managing Market Risk of Portfolio", in Barnett, W.A. and Sergi, B.S. (Eds.) *Macroeconomic Risk and Growth in the Southeast Asian Countries: Insight from SEA*, International Symposia in Economic Theory and Econometrics, Vol. 33B, Emerald Publishing Limited, Leeds, pp. 165-172, 2023.
- Gao et al., "Estimation, Inference and Empirical Analysis for Time-Varying VAR Models", *Journal of Business & Economic Statistics*, pp. 1-33, 2024.
- Helmut Lutkepohl, "Vector Autoregressive Models", Economics Working Papers ECO2011/30, European University Institute, 2011.
- Lutz Kilian, "Autorégressions vectorielles structurelles", Éditions Edward Elgar, pp. 515-554, 2013.
- Lütkepohl, H. (2004). Vector autoregressive and vector error correction models. In *New Introduction to Multiple Time Series Analysis*. doi: 10.1017/CBO9780511606885.004.
- T. Kinal and J. Ratner, "A VAR Forecasting Model of a Regional Economy: Its Construction and Comparative Accuracy", *International Regional Science Review* (SAGE Publications), Vol. 10, Iss. 2, pp. 113-126, 1986.
- Stock, J. H., & Watson, M. W. (2005). Implications of dynamic factor models for VAR analysis. *Journal of Economic Perspectives*, 19(1), 63-84.
- Stock, J. H., & Watson, M. W. (2011). Dynamic Factor Models. In *Oxford Handbook of Economic Forecasting*. doi: 10.1093/OXFORDHB/9780195398649.013.0003.