



Université Abderrahmane Mira-Bejaia

Faculté des Sciences Économiques, Commerciales et des Sciences de Gestion

Département des sciences Financières et Comptabilité

Polycopié pédagogique

Titre

Logiciels statistiques

Cours destiné aux étudiants de Master en Sciences financières et Comptabilité

Niveau : Première année

Spécialité : Comptabilité et Audit

Dr. BAAR Abdelhamid

Maître de Conférences A

Année : 2024-2025

Sommaire

Introduction générale	1
Chapitre 1 : Introduction à Eviews et régression linéaire simple	2
Chapitre 2 : La régression multiple dans Eviews	10
Chapitre 3 : Corrélation des erreurs	17
Chapitre 4 : Multiplicité linéaire (Multicolinéarité)	25
Chapitre 5 : Variance de différence (Hétéroscédasticité)	31
Chapitre 6 : Tests aléatoires	38
Chapitre 7 : Détection de la Saisonnalité et méthodes de désaisonnalisation dans Eviews	45
Chapitre 8 : Fonction d'Autocorrélation (ACF) et Fonction d'Autocorrélation Partielle (PACF)	60
Chapitre 9 : Tests de Stabilité	69
Chapitre 10 : La méthodologie de Box-Jenkins	82
Conclusion générale	90
Références bibliographiques	91
Tables des matières	93

Introduction générale

L'analyse des données économiques et financières repose sur des outils statistiques avancés permettant d'expliquer et de prédire les phénomènes observés. Parmi ces outils, la modélisation économétrique fournit un cadre rigoureux pour l'analyse des relations entre les variables. Ce cours propose une exploration approfondie des principales techniques économétriques et statistiques mises en œuvre via le logiciel Eviews, un outil simple et puissant souvent utilisé dans l'analyse des séries temporelles et la modélisation des données économiques et financières.

Le programme officiel de ce cours comporte dix axes que nous avons structurés en dix chapitres pour couvrir les concepts fondamentaux et les applications pratiques des techniques économétriques. Nous commencerons par une présentation du logiciel Eviews, afin de familiariser les étudiants avec son interface et ses fonctionnalités essentielles pour l'analyse des données. Ensuite, nous explorerons les modèles de régression linéaire simple et multiple, qui constituent la base de toute analyse économétrique.

Une bonne partie du cours sera réservée aux problèmes couramment rencontrés dans les régressions, notamment la corrélation des erreurs, la multicollinéarité et l'hétéroscédasticité. Des tests spécifiques de détection et des méthodes de correction seront abordés pour garantir la validité des estimations.

Le cours abordera également les tests de stabilité des modèles économétriques, ainsi que la détection et le traitement de la saisonnalité dans les données temporelles. Nous détaillerons l'utilisation des fonctions d'autocorrélation (ACF) et d'autocorrélation partielle (PACF), indispensables pour l'identification des structures relatives aux séries temporelles.

Enfin, le cours se conclura par une étude approfondie de la méthodologie de Box-Jenkins, qui constitue une approche méthodologique pour la modélisation et la prévision des séries temporelles à travers les modèles ARIMA.

L'objectif de ce cours est de permettre aux étudiants d'acquérir les compétences nécessaires pour mener des analyses rigoureuses et appliquées aux données économiques et financières. A travers des applications réelles dans Eviews, ils seront en mesure de comprendre les forces et les limites des outils économétriques et de les utiliser efficacement dans leurs futurs travaux académiques et professionnels.

Chapitre 1 : Introduction à Eviews et régression linéaire simple

Introduction

Dans le domaine de la comptabilité et de l'audit, l'analyse des données commerciales et financières est essentielle pour évaluer la performance des entreprises, détecter des tendances et appuyer la prise de décision. Grâce aux outils économétriques, il est possible de modéliser ces données afin d'identifier les relations entre différentes variables et d'anticiper les évolutions futures.

Parmi les logiciels les plus utilisés dans ce cadre, Eviews se distingue par sa capacité à traiter efficacement des volumes importants de données, à réaliser des analyses statistiques avancées et à estimer des modèles économétriques avec précision. Son interface simple et ses puissants outils de modélisation en font un atout majeur pour les professionnels de la comptabilité et de la finance.

Ce chapitre débutera par une présentation du logiciel Eviews, en mettant en avant ses principales fonctionnalités et son utilité dans l'analyse des données commerciales et financières. Ensuite, nous aborderons la régression linéaire, une méthode statistique fondamentale permettant d'étudier les relations entre différentes variables économiques et comptables. Nous verrons comment l'appliquer dans Eviews, interpréter les résultats obtenus et en tirer des conclusions utiles pour la gestion et l'audit des entreprises.

1) Présentation du logiciel Eviews

Eviews (Econometric Views)¹ est un logiciel d'analyse statistique et économétrique développé par Quantitative Micro Software (QMS), Il est particulièrement utilisé dans les domaines de l'économie, de la finance et des sciences sociales pour traiter des séries temporelles, effectuer des régressions, réaliser des prévisions et analyser des modèles économétriques.

1-1) Caractéristiques principales du logiciel Eviews

Le logiciel Eviews est doté de plusieurs fonctionnalités et ses caractéristiques² principales sont les suivantes :

- Interface simple : Eviews propose une interface intuitive basée sur des fenêtres et des menus déroulants, facilitant la manipulation des données et l'exécution des analyses ;

¹ Quantitative Micro Software. About Eviews. Eviews. Retrieved January 2, 2025, from https://www.eviews.com/general/about_us.html

² Quantitative Micro Software. Eviews 10 Feature List. Eviews. Retrieved January 2, 2025, from <https://www.eviews.com/EViews10/ev10features.html>

- Importation et gestion des données : Il prend en charge plusieurs formats de fichiers (Excel, CSV, Stata, SAS, etc.) et permet un traitement efficace des bases de données ;
- Outils économétriques avancés : Il permet d'estimer des modèles de régression, d'analyser des séries temporelles, d'effectuer des tests statistiques (stationnarité, autocorrélation, hétéroscédasticité...etc.), et de faire des prévisions ;
- Langage de programmation intégré : Eviews dispose d'un langage de script permettant l'automatisation des tâches et le développement de modèles économétriques complexes.

1-2) Utilisation du logiciel Eviews en analyse économique

Le logiciel Eviews est largement utilisé pour :

- L'étude de la relation entre différentes variables économiques (exemple : relation entre inflation et croissance économique) ;
- L'analyse de la performance des marchés financiers ;
- L'évaluation des politiques économiques par des modèles de prévision.

2) Régression Linéaire simple dans Eviews

2-1) Définition et principe

La régression linéaire est une méthode statistique qui permet de modéliser la relation entre une variable dépendante Y et une ou plusieurs variables indépendantes X, L'équation générale d'un modèle de régression linéaire simple est donnée par :

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Où :

- Y est la variable dépendante,
- X est la variable explicative,
- β_0 est l'ordonnée à l'origine (constante),
- β_1 est le coefficient de régression,
- ε est l'erreur aléatoire,

2-2) Estimation d'un modèle de régression linéaire simple dans Eviews

Voici les étapes pour effectuer une régression linéaire dans Eviews³ :

2-2-1) Importer ou saisir les données : Charger ou saisir un fichier contenant les variables à analyser :

³ Keil, M. W. Eviews 10 Tutorial. Retrieved January 3, 2025, from https://www.princeton.edu/~mwatson/Stock-Watson_4E/eviews_tutorial_10_1.1.pdf

a) Création d'un fichier de données dans Eviews : Pour entrer manuellement des données dans Eviews, voici les étapes à suivre :

1. Ouvrir **Eviews**.
2. Aller dans **File** → **New** → **Workfile**.
3. Une fenêtre s'affiche pour définir la structure du fichier :
 - **Données chronologiques (Time series)** : Si les données évoluent dans le temps (exemple : ventes mensuelles, bénéfices annuels),
 - **Données en coupe (Cross-section)** : Pour des observations à un instant donné (exemple : données financières de plusieurs entreprises en 2024),
 - **Données en panel** : Pour combiner séries temporelles et observations en coupe.
4. Définir la **fréquence des données** (annuelle, trimestrielle, mensuelle, journalière, personnalisée) et la période d'observation.
5. Cliquer sur **OK** pour créer le fichier de travail.
6. Dans la fenêtre du fichier de travail, cliquer sur **Objects** → **New Object** → **Series**.
7. Nommer la variable (exemple : Chiffre_Affaires).
8. Double-clique sur la série pour ouvrir une feuille de données et entrer les valeurs manuellement.

b) Importation depuis un fichier Excel ou CSV

1. **Préparation du fichier :**
 - S'assurer que la première ligne contient les **noms des variables**.
 - Vérifier que les données sont bien structurées (pas de cellules fusionnées, pas de valeurs manquantes importantes).
2. **Procédure d'importation dans Eviews :**
 - **Méthode 1 : Glisser-déposer**
 - Glisser un fichier Excel/CSV directement dans la fenêtre d'Eviews.
 - Une boîte de dialogue s'ouvre pour spécifier les paramètres d'importation.
 - Sélectionner la plage de données et cliquer sur **OK**.
 - **Méthode 2 : Menu d'importation**
 1. Aller dans **File** → **Import** → **Import from File**.
 2. Sélectionner un fichier (Excel, CSV, etc.).
 3. Spécifier les options d'importation :
 - Sélectionner la feuille contenant les données (pour Excel).

- Vérifier que la première ligne contient les noms des variables.

4. Cliquer sur **OK** pour finaliser l'importation.

2-2-2) Définir les variables : Sélectionner la variable dépendante et les variables explicatives.

2-2-3) Estimer le modèle :

- Aller dans Quick → Estimate Equation
- Entrer la spécification du modèle (ex : Y C X pour une régression simple).
- Choisir la méthode d'estimation (par défaut, Moindres Carrés Ordinaires - MCO).

Analyser les résultats :

- Examiner les coefficients β_0 et β_1 .
- Vérifier la significativité statistique (p-value, t-statistique).
- Évaluer la qualité du modèle (R^2 , tests diagnostiques).

2-3) Interprétation des résultats

Après estimation, Eviews fournit plusieurs informations utiles :

- **Les coefficients estimés** : Indiquent l'effet des variables explicatives sur la variable dépendante.
- **Le R^2** : Mesure la proportion de la variance de Y expliquée par X.
- **Les tests statistiques** : Permettent d'évaluer la validité du modèle (test de Student, test de Fisher, tests d'hypothèses sur les résidus).

La création et l'importation de fichiers de données sont des étapes essentielles pour travailler efficacement avec Eviews. Que ce soit pour entrer manuellement des données ou les importer depuis un fichier externe, Eviews offre des outils flexibles pour organiser et analyser des données financières et commerciales. Une fois cette étape réalisée, on peut facilement passer aux traitements statistiques et aux modélisations économétriques.

Exemple d'Application (Importation et analyse d'un fichier de données commerciales dans Eviews) :

Nous allons illustrer l'importation et l'analyse d'un fichier de données commerciales en utilisant Eviews. Supposons que nous disposons d'un fichier Excel (tableau 2, page 13) contenant les données annuelles d'une entreprise, avec les variables suivantes :

- Date : Période d'observation (1975 à 2024)
- Chiffre_affaires : Chiffre d'affaires annuel en millions de DA
- Coût de production: Coûts liés à la production
- Publicité : Montant des dépenses publicitaires
- Employés : Nombre d'employés

Étape 1 : Préparation des données

Avant d'importer le fichier dans Eviews, il faut s'assurer que le fichier Excel est bien structuré et ressemble au tableau 2 :

Étape 2 : Création (saisir) ou importation des données dans Eviews

1. **Ouvrir Eviews** et créer un **nouveau Workfile** :
 - Allez dans **File** → **New** → **Workfile**.
 - Sélectionner **Dated – Regular Frequency** (Séries temporelles).
 - Choisir **Annuel** et définir la période **1975 – 2024**.
 - Cliquer sur **OK**.
2. **Importer le fichier Excel** :
 - Aller dans **File** → **Import** → **Import from File**.
 - Sélectionner le fichier Excel contenant les données.
 - S'assurer que **la première ligne contient les noms des variables**.
 - Cliquer sur **OK** pour finaliser l'importation.
3. **Vérification des Données** :
 - Ouvrir le Workfile et double-clique sur les variables pour voir leur contenu.
 - Vérifier qu'aucune donnée ne manque et que les valeurs sont correctes.

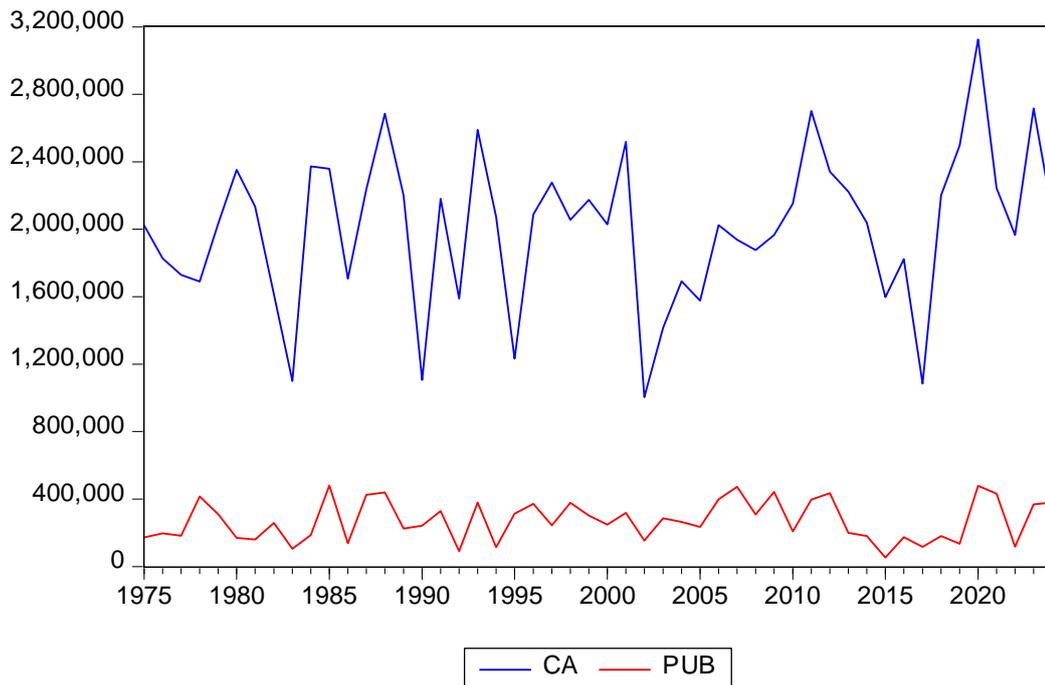
Étape 3 : Analyse des données dans Eviews

✓ *Visualisation des séries temporelles* :

Nous allons tracer l'évolution du chiffre d'affaires et des dépenses publicitaires au fil du temps :

1. Sélectionner **CA** et **PUB**.
2. Cliquer sur **View** → **Graph**.
3. Choisir **Line Graph** et cliquer sur **OK**.

Figure 1 : Représentation graphique des séries CA et PUB,



Source : Réalisée par l'auteur à partir du tableau 2 en utilisant Eviews 10.

4. Analyser le graphique (figure 1) pour identifier les tendances saisonnières et les variations.

✓ **Régression linéaire : Impact des dépenses publicitaires sur le chiffre d'affaires**

Nous allons estimer une régression linéaire pour analyser si les dépenses publicitaires influencent le chiffre d'affaires :

$$\text{Chiffre d'affaires} = \beta_0 + \beta_1 \times \text{publicité} + \varepsilon$$

Procédure dans Eviews

1. Aller dans **Quick** → **Estimate Equation**.
2. Saisir la spécification du modèle : chiffres_d'affaires c publicité
C représente la constante.
3. Cliquer sur **OK** pour exécuter l'estimation.

Tableau 1 : Estimation d'une régression linéaire simple.

Dependent Variable: CA
Method: Least Squares
Date: 01/04/25 Time: 23:06
Sample: 1975 2024
Included observations: 50

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	1530362,	143389,2	10,67278	0,0000
PUB	1,745767	0,483166	3,613184	0,0007
R-squared	0,213825	Mean dependent var		2005291,
Adjusted R-squared	0,197446	S,D, dependent var		452267,9
S,E, of regression	405165,9	Akaike info criterion		28,70116
Sum squared resid	7,88E+12	Schwarz criterion		28,77764
Log likelihood	-715,5290	Hannan-Quinn criter,		28,73028
F-statistic	13,05510	Durbin-Watson stat		1,702958
Prob(F-statistic)	0,000723			

Source : réalisé par l'auteur en utilisant Eviews 10.

Interprétation des résultats : La régression estime la relation entre le chiffre d'affaires (variable dépendante) et les dépenses publicitaires (variable explicative), L'équation obtenue est : $CA = 1\,530\,362 + 1,7458 \times PUB$. Cela signifie qu'une augmentation de 1 DA des dépenses publicitaires entraîne, en moyenne, une augmentation de 1,75 DA du chiffre d'affaires, toutes choses égales par ailleurs. Le tableau 1 affiche aussi d'autres résultats :

- Constante (C) = 1 530 362 : C'est la valeur du CA lorsque PUB = 0. Elle est significative (p-value = 0,0000 ; inférieure à 5 %), ce qui signifie qu'elle diffère significativement de zéro.
- La p-value = 0,0007 < 5 % : Le coefficient de la publicité est significatif. Donc, la publicité a un effet positif sur le CA.
- Le $R^2 = 0,2138$ (21,38 %) : Ce coefficient indique que 21,38 % de la variation du CA est expliquée par les dépenses publicitaires (PUB), Il n'est pas très élevé, ce qui suggère que d'autres variables influencent fortement le CA.

Cet exemple pratique montre comment importer un fichier de données commerciales dans Eviews, analyser des séries temporelles et estimer une régression linéaire pour examiner l'influence d'un facteur sur le chiffre d'affaires. Ces techniques sont particulièrement utiles en comptabilité, finance et audit pour l'aide à la décision et la prévision des performances d'une entreprise.

Conclusion

L'utilisation d'Eviews dans l'analyse des données commerciales et financières offre un cadre puissant pour le traitement, la modélisation et l'interprétation des relations entre différentes variables étudiées. Ce chapitre a permis d'explorer les premières étapes fondamentales de l'utilisation du logiciel, à savoir la création et l'importation d'un fichier de données, ainsi que l'application de la régression linéaire pour examiner les interactions entre variables.

Nous avons ainsi illustré comment importer et structurer une base de données, puis exploiter les outils d'Eviews pour visualiser et analyser l'évolution du chiffre d'affaires en fonction des dépenses publicitaires. Ces approches permettent aux professionnels de la comptabilité et de l'audit d'affiner leurs analyses et d'optimiser leur prise de décision sur des bases quantitatives solides.

Dans le chapitre suivant, nous approfondirons cette démarche en abordant la régression linéaire multiple avec le même exemple d'application pour permettre l'intégration des autres variables qui devraient expliquer mieux le chiffre d'affaire.

Chapitre 2 : La régression multiple dans Eviews

Introduction

La régression multiple est une extension de la régression linéaire simple qui permet d'expliquer la variable dépendante par plusieurs variables explicatives. Dans le domaine de la comptabilité et de l'audit, cette technique est particulièrement utile pour analyser l'impact de plusieurs facteurs sur des variables financières comme le chiffre d'affaires, la rentabilité ou les coûts.

Eviews, en tant que logiciel d'économétrie avancé, offre un cadre simple et efficace pour estimer des modèles de régression multiple, tester leurs performances et interpréter les résultats de manière pertinente.

1) Présentation de la régression multiple

La régression multiple suit le modèle général suivant :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

Où :

- Y est la variable dépendante (exemple : chiffre d'affaires).
- X_1, X_2, \dots, X_n sont les variables explicatives (exemple : dépenses publicitaires, coûts de production, nombre d'employés, ... etc.).
- β_0 est l'ordonnée à l'origine.
- $\beta_1, \beta_2, \dots, \beta_n$ sont les coefficients des variables explicatives.
- ε est le terme d'erreur.

L'objectif est d'estimer les coefficients pour comprendre la relation entre et ses facteurs explicatifs.

2) Estimation d'une régression multiple dans Eviews

L'estimation d'une régression multiple dans Eviews⁴ passe par plusieurs étapes.

2-1) Importation des données

Avant d'estimer un modèle de régression multiple, il est nécessaire d'importer les données dans Eviews. Cela peut se faire à partir d'un fichier Excel ou d'une base de données externe.

⁴ Samatar, E. (2023). Eviews – Econometrics Regression Analysis: A step-by-step guide. Retrieved from https://www.researchgate.net/publication/378635261_Eviews_Econometrics_Regression_Analysis_A_step-by-step_guide

1. Ouvrir Eviews et créer un nouveau fichier de travail.
2. Importer les données via **File > Import > Import from File**.
3. Vérifier l'affichage des variables dans la feuille de travail.

2-2) Lancement de la régression multiple

1. Aller dans **Quick > Estimate Equation**
2. Entrer l'équation sous la forme :

$$Y = C + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n$$
 - Y : Variable dépendante
 - C : Constante du modèle
 - $X_1 X_2 X_3 \dots X_n$: Variables explicatives
3. Choisir la méthode **LS - Least Squares (Moindres Carrés Ordinaires, MCO)**,
4. Cliquer sur **OK** pour exécuter l'estimation.

3) Interprétation des résultats

3-1) Lecture des coefficients

Chaque coefficient estimé indique l'impact de la variable explicative correspondante sur la variable dépendante. Par exemple, si le coefficient de X_1 est de 0,5, cela signifie qu'une augmentation de 1 unité de X_1 entraîne une augmentation de 0,5 unité de Y, toutes choses égales par ailleurs.

3-2) Significativité statistique

Les tests de significativité permettent de vérifier la pertinence des variables explicatives :

- **p-value (Prob)** : Si $p < 0,05$, la variable explicative est significative.
- **t-Statistic** : Permet d'évaluer l'importance relative des variables explicatives.

3-3) Qualité globale du modèle

- **R-squared (R^2)** : Indique la proportion de la variance de expliquée par le modèle.
- **Adjusted R-squared** : Réajuste en fonction du nombre de variables explicatives.
- **F-statistic** : Teste la significativité globale du modèle.

4) Test de validité du modèle

4-1) Test de multicollinéarité

La multicollinéarité survient lorsque les variables explicatives sont fortement corrélées entre elles. On peut utiliser :

- **Le facteur d'inflation de la variance (VIF)** : Si $VIF > 10$, un problème de multicollinéarité est présent.
- **La matrice de corrélation** entre les variables explicatives.

4-2) Test d'hétéroscédasticité

L'hétéroscédasticité survient lorsque la variance des erreurs n'est pas constante. On peut utiliser :

- **Test de Breusch-Pagan.**
- **Test de White.**

4-3) Test d'autocorrélation

L'autocorrélation signifie que les erreurs ne sont pas indépendantes. On utilise le **test de Durbin-Watson** (une valeur proche de 2 indique une absence d'autocorrélation).

Exemple d'application d'une régression multiple dans Eviews

Nous allons appliquer une régression multiple sur des données commerciales afin d'expliquer le chiffre d'affaires (Y) en fonction de plusieurs variables explicatives :

- Dépenses publicitaires (X_1)
- Nombre d'employés (X_2)
- Coûts de production (X_3)

Etapes de l'application

1. **Importation des données :**

- Charger un fichier Excel contenant les variables mentionnées.
- Vérifier les noms des variables et leur type.

2. **Estimation du modèle :**

- Aller dans **Quick > Estimate Equation**.
- Saisir l'équation :
CA C PUB EMPLOYES COUTS
- Choisir la méthode des **Moindres Carrés Ordinaires (MCO)** et cliquer sur **OK**.

3. **Interprétation des résultats :**

- Lire les coefficients et vérifier leur significativité (p-value < 0,05).
- Examiner le **R² ajusté** pour voir la qualité du modèle.
- Tester la multicollinéarité via le **VIF**.

4. **Tests de validité :**

- Test de Breusch-Pagan pour l'hétéroscédasticité.
- Test de Durbin-Watson pour l'autocorrélation.

Voici la base de données complète des variables de l'exemple : **CA** (Chiffre d'affaires en DA), **PUB** (Dépenses publicitaires en DA), **EMPLOYES** (Nombre d'employés) et **COUTS** (Coûts de production en DA).

Tableau 2 : Base de données de la régression linéaire simple et multiple.

Année	CA	PUB	EMPLOYES	COUTS
1975	2022154	171958	73	1272876
1976	1826703	196867	12	1715857
1977	1729753	181932	60	807086
1978	1689745	415838	16	733556
1979	2031787	309178	30	1597527
1980	2352916	169879	82	1522905
1981	2135186	160268	48	1967640
1982	1613867	257892	27	942139
1983	1100496	104886	13	632315
1984	2372102	187337	98	1126075
1985	2358623	480410	69	716588
1986	1707861	137498	23	1485160
1987	2237643	424871	18	1810908
1988	2685662	438468	99	1178732
1989	2199611	225203	62	1625472
1990	1106011	241335	11	248984
1991	2180497	328167	93	502918
1992	1589395	91090	69	464712
1993	2588660	379365	80	1368727
1994	2072468	114820	53	1704384
1995	1232905	312913	17	333767
1996	2086781	371879	56	848663
1997	2276841	244027	44	1952126
1998	2056001	377069	87	336330
1999	2173702	302709	90	680754
2000	2028736	249041	45	1550224
2001	2517526	317455	59	1937095
2002	1004565	153355	13	213986
2003	1415386	285796	11	786146
2004	1692128	264176	15	1377888
2005	1575592	234779	63	212666
2006	2023911	397449	13	1549380
2007	1937711	471909	63	334633
2008	1876738	308795	72	488998
2009	1965947	442752	27	1050937
2010	2152011	206730	99	828776
2011	2700282	396622	53	1956187
2012	2341046	434681	43	1533382
2013	2223461	199503	83	1279111

2014	2037970	180523	71	1196107
2015	1597109	52747	23	1624289
2016	1821635	173855	57	1038688
2017	1084647	115725	24	551279
2018	2200114	179981	81	1113910
2019	2496325	134654	87	1850237
2020	3125389	478989	96	1896548
2021	2242914	431490	71	689570
2022	1965986	117435	49	1705127
2023	2716471	369030	94	1521685
2024	2093560	378947	89	224538

Source : données fictives.

Cet exemple met en évidence l'utilisation pratique d'une régression multiple dans Eviews. En appliquant les tests et en interprétant les résultats, nous pouvons déterminer les facteurs influençant le chiffre d'affaires et leur impact sur la performance financière.

Résultats de la régression multiple : L'estimation du modèle de régression multiple donne les résultats suivants :

Tableau 3 : Résultats de la régression linéaire multiple.

Dependent Variable: CA
Method: Least Squares
Date: 02/01/25 Time: 21:30
Sample: 1975 2024
Included observations: 50

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	467028,1	24527,27	19,04117	0,0000
PUB	1,592025	0,056514	28,17061	0,0000
EMPLOYES	9871,890	230,6568	42,79903	0,0000
COUTS	0,504529	0,012002	42,03743	0,0000
R-squared	0,989970	Mean dependent var		2005291,
Adjusted R-squared	0,989316	S,D, dependent var		452267,9
S,E, of regression	46747,79	Akaike info criterion		24,41954
Sum squared resid	1,01E+11	Schwarz criterion		24,57250
Log likelihood	-606,4885	Hannan-Quinn criter,		24,47779
F-statistic	1513,444	Durbin-Watson stat		2,059957
Prob(F-statistic)	0,000000			

Source : réalisé par l'auteur en utilisant Eviews 10.

Equation estimée :

$$\widehat{CA} = 467000 + 1,592 \times PUB + 9871,89 \times EMPLOYES + 0,5045 \times COUTS$$

Interprétation des coefficients :

- **Constante ($4,67 \times 10^5$) :** Lorsque toutes les variables explicatives sont nulles, le chiffre d'affaires est estimé à 467 000 DA.

- **PUB (1,592)** : Une augmentation de 1 DA des dépenses publicitaires entraîne une augmentation moyenne de 1,592 DA du chiffre d'affaires, toutes choses égales par ailleurs.
- **EMPLOYES (9871,89)** : Chaque employé supplémentaire génère en moyenne une augmentation de 9871,89 DA du chiffre d'affaires.
- **COUTS (0,5045)** : Une augmentation de 1 DA des coûts de production entraîne une augmentation moyenne de 0,5045 DA du chiffre d'affaires.

Statistiques clés :

- **R² ajusté = 0,989** : Le modèle explique 98,9% de la variation du chiffre d'affaires,
- **p-values < 0,05 pour toutes les variables** : Tous les coefficients sont statistiquement significatifs.
- **Durbin-Watson = 2,060** : Pas d'autocorrélation significative des erreurs.

Le modèle est bien ajusté avec des coefficients significatifs et une bonne capacité explicative (R² élevé). Cependant, les erreurs standard élevées 46747,79 pourraient indiquer une collinéarité entre les variables explicatives, Une analyse plus approfondie pourrait être nécessaire (ex: test du Facteur d'inflation de la variance "**Variance Inflation Factor**" (VIF)). Nous allons détailler ce test de multicollinéarité dans le quatrième chapitre.

La procédure du test VIF dans Eviews⁵ est la suivante :

1. **Ouvrir le modèle estimé** : Après avoir effectué la régression multiple, sélectionne l'équation estimée dans la fenêtre de travail.
2. **Tester la multicollinéarité avec le VIF** :
 - Aller dans **View > Coefficient Diagnostics > Variance Inflation Factors (VIF)**.
 - Eviews affichera les valeurs du VIF pour chaque variable explicative.

Interprétation des résultats :

- **VIF < 5** → Pas de problème de collinéarité.
- **5 ≤ VIF < 10** → Colinéarité modérée, à surveiller.
- **VIF ≥ 10** → Forte colinéarité, il faut peut-être exclure ou transformer certaines variables.

⁵ IHS Global Inc. (2023). Eviews 12 User's Guide. IHS Global Inc. Retrieved from <https://www.eviews.com>

Conclusion

La régression multiple est un outil puissant pour analyser les relations entre plusieurs variables dans un cadre commercial et financier. Eviews permet une estimation rapide et efficace des modèles de régression multiple, tout en offrant divers tests de validité pour assurer la pertinence des résultats.

Dans le chapitre suivant, nous approfondirons l'analyse économétrique en développant la notion et les tests de corrélation des erreurs.

Chapitre 3 : Corrélation des erreurs

Introduction

Dans une régression linéaire multiple, l'hypothèse d'absence de corrélation des erreurs (autocorrélation) est essentielle pour assurer l'efficacité des estimateurs des moindres carrés ordinaires (MCO). Si les erreurs sont corrélées, les estimateurs restent sans biais, mais deviennent inefficaces, ce qui affecte la précision des prédictions et des inférences statistiques⁶. Ce chapitre explore les tests permettant de détecter la corrélation des erreurs et les méthodes pour y remédier dans Eviews.

1) Tests de détection de la corrélation des erreurs

1-1) Test de Durbin-Watson

Le test de Durbin-Watson est l'un des tests les plus utilisés pour détecter l'autocorrélation des erreurs, en particulier dans les séries temporelles. Il est basé sur la statistique suivante :

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

Interprétation :

- Si $DW \approx 2$, il n'y a pas d'autocorrélation.
- Si $DW < 2$, il y a une autocorrélation positive.
- Si $DW > 2$, il y a une autocorrélation négative.

Dans Eviews, la statistique Durbin-Watson est affichée automatiquement après l'estimation d'un modèle de régression.

1-2) Test de Breusch-Godfrey

Le test de Breusch-Godfrey est une alternative plus générale au test de Durbin-Watson, permettant de détecter une autocorrélation d'ordre supérieur.

Hypothèses du test :

- H_0 : Absence d'autocorrélation.
- H_1 : Présence d'autocorrélation.

Dans Eviews, ce test peut être effectué en sélectionnant l'option « Residual Diagnostics » après l'estimation d'une régression.

1-3) Test de Ljung-Box

Le test de Ljung-Box vérifie si les résidus présentent une autocorrélation significative à plusieurs retards.

⁶ Gujarati, D. N., & Porter, D. C. (2021). Basic Econometrics (5th ed.). McGraw-Hill.

Hypothèses du test :

- H_0 : Les erreurs ne sont pas corrélées.
- H_1 : Les erreurs sont corrélées.

Ce test est particulièrement utile pour les modèles de séries temporelles et peut être effectué via l'analyse des résidus dans Eviews.

2) Résolution des problèmes de corrélation des erreurs

2-1) Transformation de Cochrane-Orcutt

La méthode de Cochrane-Orcutt consiste à transformer les données pour éliminer l'autocorrélation des erreurs. Elle repose sur le modèle : $e_t = \rho e_{t-1} + u_t$

Si ρ est estimé, on peut appliquer une transformation pour neutraliser l'autocorrélation.

Dans Eviews, cette correction est disponible via l'option de régression en « GLS ».

2-2) Modélisation avec les erreurs ARMA

Lorsque l'autocorrélation est complexe, une modélisation des erreurs sous forme de processus ARMA (AutoRegressive Moving Average) peut être nécessaire. On intègre alors les erreurs dans un modèle du type : $e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \dots + u_t$

Eviews permet d'inclure ces termes dans la spécification du modèle. Ce point sera détaillé dans le huitième chapitre.

2-3) Utilisation des moindres carrés généralisés (GLS)

La méthode GLS ajuste les moindres carrés ordinaires en tenant compte de la structure d'autocorrélation, rendant les estimateurs plus efficaces. Cette approche est souvent privilégiée lorsque l'autocorrélation est confirmée.

Exemple d'application (Détection et correction de l'autocorrélation dans une régression financière) :

Un analyste financier modélise la relation entre le prix du pétrole (\$) et l'indice boursier S&P 500. En utilisant les données mensuelles de janvier 2021 à décembre 2023 (tableau 8, page 24), il obtient la régression suivante :

$$SP500_t = \beta_0 + \beta_1 \times Prix_Petrole_t + \varepsilon_t$$
$$SP500_t = 2479,16 + 23,50 \times Prix_Petrole_t + \varepsilon_t$$

Tableau 4 : Régression de l'indice boursier sur les prix du pétrole.

Dependent Variable: SP500
Method: Least Squares
Date: 02/11/25 Time: 10:54
Sample: 2021M01 2023M12
Included observations: 36

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	2479,155	64,92049	38,18755	0,0000
PRIX_PETROLE	23,49624	0,712234	32,98950	0,0000
R-squared	0,969705	Mean dependent var		4577,500
Adjusted R-squared	0,968814	S,D, dependent var		441,5452
S,E, of regression	77,97462	Akaike info criterion		11,60460
Sum squared resid	206721,4	Schwarz criterion		11,69257
Log likelihood	-206,8827	Hannan-Quinn criter,		11,63530
F-statistic	1088,307	Durbin-Watson stat		0,302869
Prob(F-statistic)	0,000000			

Source : réalisé par l'auteur en utilisant Eviews 10.

Interprétation des coefficients :

- Le coefficient constant ($\beta_0 = 2479,16$) signifie que lorsque le prix du pétrole est nul, l'indice S&P 500 est estimé à 2479,16.
- Le coefficient associé au Prix_Petrole ($\beta_1 = 23,50$) indique qu'une augmentation de 1 \$ du prix du pétrole entraîne une hausse de 23,50 points de l'indice S&P 500, en moyenne.

Statistiques globales :

- $R^2 = 0,97$: Le modèle explique 97 % de la variance de l'indice S&P 500, ce qui montre une forte adéquation du modèle aux données.
- F-statistique = 1088,31 (p-valeur = 0,0000) : Le modèle est globalement significatif au seuil de 1 %.

La statistique Durbin-Watson ($DW = 0,30$) est bien inférieure à 2, indiquant une autocorrélation positive importante des résidus. Cela peut biaiser les écarts-types et fausser les tests de significativité. Pour corriger cette autocorrélation détectée dans la régression, trois (3) méthodes sont possibles :

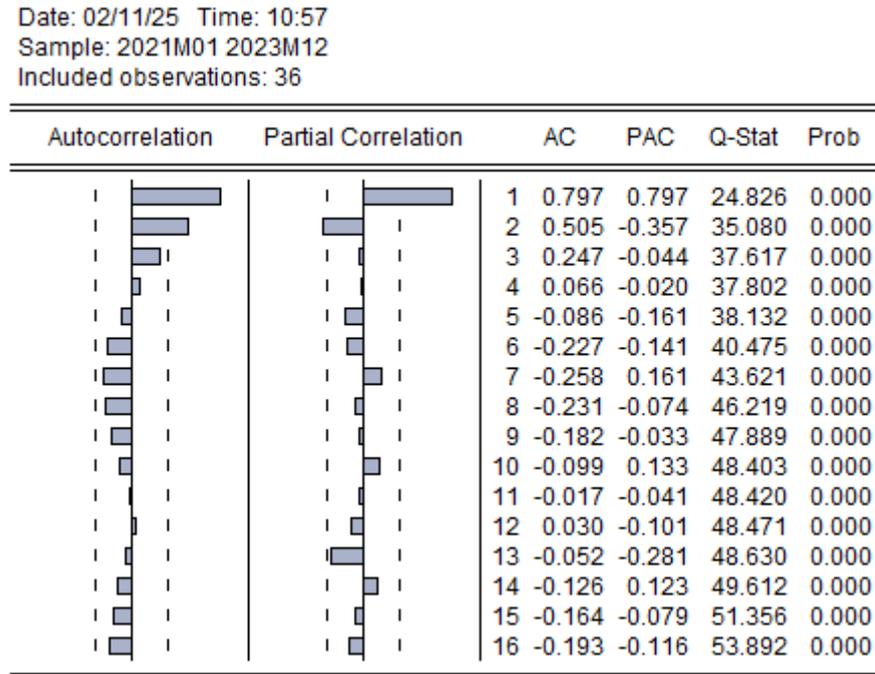
- ✓ **Méthode de Cochrane-Orcutt** : Cette méthode repose sur une transformation de la régression pour éliminer l'autocorrélation.

Étapes sous Eviews :

1. Aller dans **Quick** → **Estimate Equation** et entrez votre équation.

2. Une fois la régression estimée, cliquer sur **View** → **Residual Diagnostics** → **Correlogram – Q-statistics** pour confirmer l'autocorrélation.

Figure 2 : Correlogramme des résidus.



Source : réalisée par l'auteur en utilisant Eviews 10.

3. Pour appliquer la correction :

- Aller dans **Proc** → **Make Residual Series** pour créer une série des résidus.
- Estimer le modèle AR(1) en ajoutant un terme autorégressif à votre équation :

$$SP500_t = \beta_0 + \beta_1 \times Prix_Petrole_t + \rho \times e_{t-1} + u_t$$

Tableau 5 : Résultats de la régression de l'indice boursier sur les prix du pétrole (Cochrane-Orcutt).

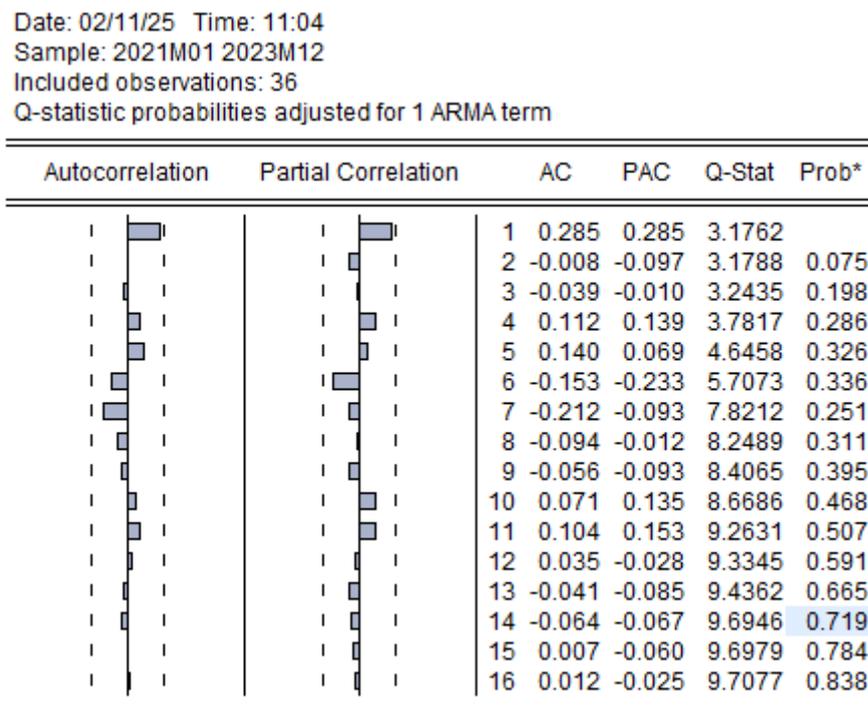
Dependent Variable : SP500
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/11/25 Time: 11:02
Sample: 2021M01 2023M12
Included observations: 36
Convergence achieved after 97 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	2761,675	640,5773	4,311229	0,0001
PRIX_PETROLE	20,53804	5,507801	3,728901	0,0007

AR(1)	0,896441	0,107152	8,366092	0,0000
SIGMASQ	1517,699	331,4490	4,578982	0,0001
R-squared	0,991993	Mean dependent var		4577,500
Adjusted R-squared	0,991242	S,D, dependent var		441,5452
S,E, of regression	41,32083	Akaike info criterion		10,43026
Sum squared resid	54637,16	Schwarz criterion		10,60621
Log likelihood	-183,7447	Hannan-Quinn criter,		10,49167
F-statistic	1321,501	Durbin-Watson stat		1,420913
Prob(F-statistic)	0,000000			
Inverted AR Roots	,90			

Source : réalisé par l'auteur en utilisant Eviews 10.

Figure 3 : Corrélogramme des résidus.



*Probabilities may not be valid for this equation specification.

Source : réalisée par l'auteur en utilisant Eviews 10.

- ✓ **Modélisation des erreurs ARMA** : Si l'autocorrélation est complexe, les erreurs sont modélisées comme un processus ARMA.

Étapes sous Eviews :

1. Aller dans **Quick** → **Estimate Equation**.
2. Spécifier l'équation avec les termes AR(p) et MA(q), par exemple :

$$SP500_t = \beta_0 + \beta_1 \times \text{Prix_Petrole}_t + AR(1) + MA(1)$$

3. Vérifier les nouvelles statistiques Durbin-Watson et les tests de Breusch-Godfrey.

Tableau 6 : Résultats de la régression de l'indice boursier sur les prix du pétrole (ARMA).

Dependent Variable: SP500
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/11/25 Time: 11:08
Sample: 2021M01 2023M12
Included observations: 36
Convergence achieved after 80 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	2652,488	197,5554	13,42655	0,0000
PRIX_PETROLE	21,67376	2,318818	9,346899	0,0000
AR(1)	0,799617	0,167899	4,762485	0,0000
MA(1)	0,352210	0,192874	1,826115	0,0775
SIGMASQ	1375,416	314,0550	4,379538	0,0001

R-squared	0,992744	Mean dependent var	4577,500
Adjusted R-squared	0,991807	S,D, dependent var	441,5452
S,E, of regression	39,96570	Akaike info criterion	10,38796
Sum squared resid	49514,98	Schwarz criterion	10,60790
Log likelihood	-181,9833	Hannan-Quinn criter,	10,46473
F-statistic	1060,280	Durbin-Watson stat	1,879133
Prob(F-statistic)	0,000000		

Inverted AR Roots	,80
Inverted MA Roots	-,35

Source : réalisé par l'auteur en utilisant Eviews 10.

Figure 4 : Corrélogramme des résidus.

Date: 02/11/25 Time: 11:09
Sample: 2021M01 2023M12
Included observations: 36
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*	
		1	0.044	0.044	0.0761	
		2	0.048	0.046	0.1695	
		3	-0.058	-0.063	0.3105	0.577
		4	0.074	0.078	0.5446	0.762
		5	0.151	0.152	1.5466	0.672
		6	-0.157	-0.188	2.6654	0.615
		7	-0.131	-0.128	3.4711	0.628
		8	-0.016	0.037	3.4842	0.746
		9	-0.067	-0.105	3.7133	0.812
		10	0.073	0.068	3.9975	0.857
		11	0.090	0.190	4.4431	0.880
		12	0.082	0.063	4.8287	0.902
		13	-0.054	-0.121	5.0001	0.931
		14	-0.077	-0.068	5.3697	0.944
		15	-0.004	-0.055	5.3709	0.966
		16	-0.005	-0.074	5.3727	0.980

*Probabilities may not be valid for this equation specification.

Source : réalisée par l'auteur en utilisant Eviews 10.

- ✓ **Utilisation des moindres carrés généralisés (GLS) :** Le GLS ajuste la régression pour corriger l'autocorrélation automatiquement.

Étapes sous Eviews :

1. Aller dans **Quick** → **Estimate Equation**.
2. Sous **Estimation Settings**, sélectionner **GLS** dans la méthode d'estimation.
3. Choisir **AR(1)** dans la case « Autoregressive error specification ». Cela ajoutera un terme AR(1) pour corriger l'autocorrélation.

Tableau 7 : Résultats de la régression de l'indice boursier sur les prix du pétrole (GLS).

Dependent Variable: SP500
Method: Least Squares
Date: 02/11/25 Time: 19:56
Sample: 2021M01 2023M12
Included observations: 36
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4,0000)

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	2479,155	150,6907	16,45195	0,0000
PRIX_PETROLE	23,49624	1,498707	15,67768	0,0000
R-squared	0,969705	Mean dependent var		4577,500
Adjusted R-squared	0,968814	S,D, dependent var		441,5452
S,E, of regression	77,97462	Akaike info criterion		11,60460
Sum squared resid	206721,4	Schwarz criterion		11,69257
Log likelihood	-206,8827	Hannan-Quinn criter,		11,63530
F-statistic	1088,307	Durbin-Watson stat		0,302869
Prob(F-statistic)	0,000000	Wald F-statistic		245,7896
Prob(Wald F-statistic)	0,000000			

Source : réalisé par l'auteur en utilisant Eviews 10.

La **méthode des moindres carrés généralisés (GLS)** dans **Eviews** est disponible, mais elle n'est pas directement nommée comme "GLS". Elle est intégrée sous forme de corrections pour l'autocorrélation et l'hétéroscédasticité. D'où, cette régression pondérée équivalente à GLS pour corriger les coefficients covariance et obtenir des écarts-types robustes (HAC). Cependant, la statistique de Durbin-Watson plaide toujours pour l'existence d'une autocorrélation des erreurs.

Tableau 8 : Données sur l'indice boursier et les prix du pétrole.

Mois	Prix_Petrole	SP500	Mois	Prix_Petrole	SP500
janv-21	50	3800	juil-22	99	4780
févr-21	52	3850	août-22	102	4820
mars-21	55	3900	sept-22	100	4800
avr-21	60	3950	oct-22	98	4780
mai-21	65	4000	nov-22	95	4740
juin-21	70	4050	déc-22	92	4700
juil-21	75	4100	janv-23	94	4750
août-21	73	4080	févr-23	96	4780
sept-21	72	4070	mars-23	100	4850
oct-21	74	4090	avr-23	102	4900
nov-21	78	4150	mai-23	104	4950
déc-21	80	4200	juin-23	106	5000
janv-22	85	4500	juil-23	108	5050
févr-22	87	4550	août-23	110	5100
mars-22	90	4600	sept-23	112	5150
avr-22	92	4630	oct-23	114	5200
mai-22	95	4700	nov-23	115	5220
juin-22	97	4750	déc-23	118	5250

Source : données fictives.

Conclusion

L'autocorrélation des erreurs peut biaiser les résultats d'une régression et fausser les inférences statistiques. Il est donc nécessaire de la détecter et de la corriger. Les tests de Durbin-Watson, Breusch-Godfrey et Ljung-Box permettent d'identifier ce problème, tandis que des méthodes comme la transformation de Cochrane-Orcutt, l'ajout de termes ARMA et les moindres carrés généralisés offrent des solutions adaptées. Eviews fournit des outils performants pour réaliser ces analyses et améliorer la qualité des modèles économétriques.

Chapitre 4 : Multiplicité linéaire (Multicolinéarité)

Introduction

Dans une régression linéaire multiple, l'hypothèse d'indépendance linéaire entre les variables explicatives est essentielle. Lorsqu'il existe une forte corrélation linéaire entre ces variables, on parle de multiplicité linéaire ou multicollinéarité⁷. Ce phénomène peut rendre les estimateurs instables et diminuer la précision des prédictions. Ce chapitre présente les tests permettant de détecter la multicollinéarité ainsi que les solutions pour y remédier.

1) Détection de la multiplicité linéaire

1-1) Analyse des corrélations croisées

La première étape consiste à examiner la matrice des corrélations entre les variables explicatives. Une forte corrélation (supérieure à 0,8 ou inférieure à -0,8) peut indiquer la présence de multicollinéarité.

1-1-1) Facteur d'inflation de la variance (Variance Inflation Factor – VIF)

Le VIF mesure l'ampleur de la multicollinéarité pour chaque variable explicative. Il est calculé comme suit :

$$VIF_i = \frac{1}{1 - R_i^2}$$

Où R_i^2 est le coefficient de détermination obtenu en régressant la variable X_i sur les autres variables explicatives.

Pour accéder aux tests VIF dans Eviews, après l'estimation du modèle, il faut cliquer sur **View** dans la fenêtre des résultats, Puis, sur **Coefficient Diagnostics** → **Variance Inflation Factors**.

Interprétation :

- VIF=1: Absence de multicollinéarité.
- VIF>5 : Multicollinéarité modérée.
- VIF>10 : Multicollinéarité sévère, nécessitant une correction.

⁷ Wooldridge, J. M. (2020). Introductory Econometrics: A Modern Approach (7th ed.). Cengage Learning.

1-1-2) Test de Farrar-Glauber

Le test de Farrar-Glauber est une méthode statistique qui repose sur la matrice de corrélation pour détecter la multicolinéarité à travers plusieurs sous-tests (Chi-carré, F-test, .. etc.).

Étapes pour le test de Farrar-Glauber dans Eviews :

1. Estimer le modèle de régression.
2. Afficher la matrice de corrélation des variables explicatives :
 - Sélectionner toutes les variables explicatives.
 - Cliquer sur **View** → **Covariance Analysis** → **Correlation Matrix**.
 - Observer les coefficients de corrélation : des valeurs proches de $\pm 0,8$ peuvent indiquer une multicolinéarité.
3. **Calculer les statistiques de Farrar-Glauber:**

Le test de Farrar-Glauber se divise en trois sous-tests :

Test du Chi-carré global : Il mesure la présence globale de multicolinéarité dans le modèle. La statistique de ce test est la suivante :

$$\chi^2 = - \left(n - 1 - \frac{2p+5}{6} \right) \times \ln(|R|)$$

- n : nombre d'observations
- p : nombre de variables explicatives
- |R| : déterminant de la matrice de corrélation

Tests F individuels : Chaque variable explicative est régressée sur les autres variables, et la statistique F est utilisée pour détecter la multicolinéarité individuelle.

Test de détermination conditionnelle : Il analyse les valeurs propres de la matrice de corrélation pour identifier la sévérité de la multicolinéarité.

Interprétation des résultats :

- Si le **test global Chi-carré** est significatif, cela indique une multicolinéarité dans le modèle.
- Les **tests F individuels** permettent de localiser les variables problématiques.
- Le **test de détermination conditionnelle** identifie si la multicolinéarité est faible, modérée ou sévère.

Le test de Farrar-Glauber ne peut se faire directement sous Eviews car le calcul du déterminant de la matrice de corrélation et les statistiques associées doit se faire manuellement ou en utilisant des logiciels complémentaires (R, Python ou Excel).

1-2) Analyse des coefficients instables

La multicolinéarité peut également se détecter lorsque les signes ou les significations des coefficients changent de manière importante lorsque l'on ajoute ou supprime une variable du modèle. Une fois la multicolinéarité du modèle de base estimé est vérifiée, les coefficients obtenus ainsi que leur signe et leur significativité sont notés pour les comparer au modèle réduit (après suppression d'une variable). Il s'agit précisément d'observer les changements dans les coefficients restants :

- Magnitude : Les valeurs des coefficients changent-elles fortement ?
- Signe : Les coefficients changent-ils de signe (exemple : de positif à négatif) ?
- Significativité : Un coefficient devient-il significatif ou insignifiant après la modification du modèle ?

En effet, une grande instabilité de ces coefficients est un signe de multicolinéarité.

2) Résolution des problèmes de multiplicité linéaire

2-1) Élimination de variables explicatives

Une solution simple consiste à supprimer l'une des variables fortement corrélées. Cette méthode peut toutefois entraîner une perte d'information importante.

2-2) Transformation des variables

Certaines transformations (différences premières, logarithmes) peuvent réduire la multicolinéarité en modifiant la relation linéaire entre les variables.

2-3) Combinaison linéaire des variables

Il est possible de créer des indices composites ou des variables synthétiques qui regroupent les informations de plusieurs variables.

2-4) Régression Ridge (Ridge Regression)

Cette méthode introduit une pénalisation qui réduit l'influence de la multicolinéarité. Contrairement aux moindres carrés ordinaires (OLS), la régression Ridge minimise :

$$MIN \left(\sum (y_i - X\beta_j)^2 + \lambda \sum \beta_j^2 \right)$$

Avec λ étant le paramètre de régularisation.

Exemple d'application : Analyse de la relation entre les ventes, les dépenses publicitaires et le revenu moyen (données du tableau 9), Un analyste commercial cherche à expliquer les ventes mensuelles d'un produit en fonction des dépenses publicitaires (milliers de DA) et du revenu moyen (milliers de DA) sur la période de janvier 2021 à décembre 2023. Les données sont les suivantes :

Tableau 9 : Données sur les ventes, dépenses publicitaires et le revenu moyen.

Mois	Depenses_Pub	Revenu_Moyen	Ventes
Jan-21	1000	4000	500
Fév-21	1200	4100	530
Mar-21	1500	4200	600
Avr-21	1600	4300	620
Mai-21	1700	4400	640
Juin-21	1800	4500	660
Juil-21	1900	4600	680
Août-21	2000	4700	700
Sep-21	2100	4800	720
Oct-21	2200	4900	740
Nov-21	2300	5000	760
Déc-21	2400	5100	780
Jan-22	1100	4200	520
Fév-22	1300	4300	550
Mar-22	1600	4400	620
Avr-22	1700	4500	640
Mai-22	1800	4600	660
Juin-22	1900	4700	680
Juil-22	2000	4800	700
Août-22	2100	4900	720
Sep-22	2200	5000	740
Oct-22	2300	5100	760
Nov-22	2400	5200	780
Déc-22	2500	5300	800
Jan-23	1200	4400	540
Fév-23	1400	4500	570
Mar-23	1700	4600	640
Avr-23	1800	4700	660
Mai-23	1900	4800	680
Juin-23	2000	4900	700
Juil-23	2100	5000	720
Août-23	2200	5100	740
Sep-23	2300	5200	760

Oct-23	2400	5300	780
Nov-23	2500	5400	800
Déc-23	3000	5500	820

Source : données fictives.

Étapes dans Eviews :

1. Importation des données dans Eviews,
2. Estimation d'une régression linéaire :

$$Ventes_t = \beta_0 + \beta_1 \times Dépenses_Pub_t + \beta_2 \times Revenu_Moyen_t + \varepsilon_t$$

Tableau 10 : Régression des ventes sur les dépenses publicitaires et le revenu moyen.

Dependent Variable: VENTES
Method: Least Squares
Date: 02/11/25 Time: 14:02
Sample: 2021M01 2023M12
Included observations: 36

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	250,0746	61,76986	4,048490	0,0003
DEPENSES_PUB	0,173528	0,016549	10,48585	0,0000
REVENU_MOYEN	0,020564	0,019240	1,068837	0,2929
R-squared	0,979003	Mean dependent var		680,8333
Adjusted R-squared	0,977731	S,D, dependent var		86,69734
S,E, of regression	12,93776	Akaike info criterion		8,037833
Sum squared resid	5523,726	Schwarz criterion		8,169793
Log likelihood	-141,6810	Hannan-Quinn criter,		8,083890
F-statistic	769,3351	Durbin-Watson stat		1,112562
Prob(F-statistic)	0,000000			

Source : réalisé par l'auteur en utilisant Eviews 10,

Vérification de la multicolinéarité à l'aide du VIF : Si $VIF > 10$ pour une ou plusieurs variables, une correction est nécessaire.

Tableau 11 : Résultat du test VIF.

Variance Inflation Factors
Date: 02/11/25 Time: 14:11
Sample: 2021M01 2023M12
Included observations: 36

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	3815,515	820,6114	NA
DEPENSES_PUB	0,000274	228,6147	11,61052
REVENU_MOYEN	0,000370	1807,924	11,61052

Source : réalisé par l'auteur en utilisant Eviews 10.

Correction :

- Les dépenses publicitaires et le revenu moyen sont fortement corrélés ($VIF > 10$), on peut éliminer l'une des variables ou appliquer une transformation en différences premières.
- On peut également utiliser une régression Ridge pour réduire l'impact de la multicollinéarité.

Tableau 12 : Résultats après correction (par élimination de la variable "Revenu Moyen").

Variance Inflation Factors
Date: 02/11/25 Time: 14:17
Sample: 2021M01 2023M12
Included observations: 36

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	91,93549	19,69030	NA
DEPENSES_PUB	2,37E-05	19,69030	1,000000

Source : réalisé par l'auteur en utilisant Eviews 10.

D'où l'absence de multicollinéarité.

Conclusion

La multiplicité linéaire peut compliquer l'analyse des régressions et réduire la fiabilité des résultats. Les tests comme le VIF, l'analyse des corrélations et le test de Farrar-Glauber permettent de la détecter. Des solutions telles que l'élimination de variables, la transformation de données ou la régularisation (Ridge Regression) permettent d'améliorer les modèles économétriques. Eviews fournit des outils performants pour ces analyses, facilitant ainsi la prise de décision basée sur les données.

Chapitre 5 : Variance de différence (Hétéroscédasticité)

Introduction

L'hypothèse d'homoscédasticité (variance constante des erreurs) est essentielle dans la régression linéaire. Lorsqu'elle est violée, on parle d'hétéroscédasticité, ce qui peut entraîner des erreurs standards biaisées, rendant les tests de significativité incorrects. Ce chapitre présente les tests permettant de détecter la variance de différence et les méthodes pour y remédier.

1) Détection de l'hétéroscédasticité

1-1) Analyse graphique

La détection peut commencer par une analyse graphique. Le graphique des résidus en fonction des valeurs prédites analysé permet de visualiser la structure, ainsi une structure en "entonnoir" c'est-à-dire une variance des résidus qui augmente ou diminue est un signe d'hétéroscédasticité.

1-2) Test de Breusch-Pagan

Le test de Breusch-Pagan⁸ repose sur la régression des résidus au carré (e_i^2) sur les variables explicatives pour détecter l'hétéroscédasticité.

Hypothèses :

- H_0 : Homoscédasticité (variance constante)
- H_1 : Hétéroscédasticité

Une statistique de test élevée (ou une p-valeur faible) indique la présence d'hétéroscédasticité.

1-3) Test de White

Le test de White⁹ est une version plus générale qui ne nécessite aucune hypothèse spécifique sur la forme de l'hétéroscédasticité. Il consiste à regresser les résidus au carré sur les variables explicatives ainsi que leurs combinaisons quadratiques et croisées.

1-4) Test de Goldfeld-Quandt

Le test de Goldfeld-Quandt¹⁰ consiste à diviser les observations en deux sous-groupes, puis à comparer les variances des résidus dans chaque groupe.

⁸ Breusch, T. S. & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 47(5), 1287–1294.

⁹ White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4), 817–838.

¹⁰ Goldfeld, S. M., & Quandt, R. E. (1965). Some Tests for Homoscedasticity. *Journal of the American Statistical Association*, 60(310), 539-547.

Hypothèses :

- H_0 : Homoscédasticité
- H_1 : Hétéroscédasticité

Une forte différence entre les variances des deux groupes suggère une hétéroscédasticité.

2) Résolution des problèmes de variance de différence

2-1) Transformation des données

Une des solutions les plus simples est de transformer les données pour stabiliser la variance :

- **Transformation logarithmique** : Appliquer $\ln(Y)$ sur la variable dépendante pour réduire l'hétéroscédasticité.
- **Différences premières** : Utile lorsque les données montrent une tendance ou une croissance exponentielle.

2-2) Modèle des moindres carrés pondérés (WLS)

Le modèle des moindres carrés pondérés corrige l'hétéroscédasticité en attribuant un poids à chaque observation en fonction de l'inverse de la variance estimée :

$$w_i = \frac{1}{\hat{\sigma}_i^2}$$

Cela donne plus de poids aux observations avec une petite variance.

2,3) Utilisation des erreurs standards robustes (White)

Les erreurs standards robustes corrigent les erreurs de la variance non constante sans modifier les coefficients de régression. Elles permettent des tests statistiques fiables malgré l'hétéroscédasticité.

Exemple d'application sous Eviews : Analyse de la relation entre les ventes mensuelles d'un produit en fonction des dépenses publicitaires et du prix de vente sur une période donnée (données du tableau 18, pages 33). Les données montrent une possible variance de différence (plus les dépenses augmentent, plus les résidus semblent dispersés).

Étapes dans Eviews :

1. Estimation d'un modèle de régression linéaire :

$$\text{SALES}_t = \beta_0 + \beta_1 \times \text{ADVERTISING_EXPENSES}_t + \beta_2 \times \text{PRICE}_t + \varepsilon_t$$

Tableau 13 : Régression des ventes sur les dépenses publicitaires et le prix de vente.

Dependent Variable: SALES
 Method: Least Squares
 Date: 02/12/25 Time: 20:45
 Sample: 2021M01 2023M11
 Included observations: 35

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	967,5902	293,0216	3,302112	0,0024
ADVERTISING_EXPENSES	0,845301	0,067466	12,52928	0,0000
PRICE	-21,03036	7,355493	-2,859137	0,0074
R-squared	0,839445	Mean dependent var		2505,676
Adjusted R-squared	0,829411	S,D, dependent var		1289,897
S,E, of regression	532,7595	Akaike info criterion		15,47583
Sum squared resid	9082647,	Schwarz criterion		15,60915
Log likelihood	-267,8271	Hannan-Quinn criter,		15,52185
F-statistic	83,65449	Durbin-Watson stat		2,167534
Prob(F-statistic)	0,000000			

Source : réalisé par l'auteur en utilisant Eviews 10.

2. Test de Breusch-Pagan : Si $p < 0,05$ il y a présence d'hétéroscédasticité.

Tableau 14 : Résultat du test de Breusch-Pagan.

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	4,466142	Prob, F(2,32)	0,0195
Obs*R-squared	7,637735	Prob, Chi-Square(2)	0,0220
Scaled explained SS	7,452572	Prob, Chi-Square(2)	0,0241

Test Equation:
 Dependent Variable: RESID^2
 Method: Least Squares
 Date: 02/12/25 Time: 20:46
 Sample: 2021M01 2023M11
 Included observations: 35

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	-75853,44	201659,0	-0,376147	0,7093
ADVERTISING_EXPENSES	138,4544	46,43048	2,981971	0,0054
PRICE	-588,3865	5062,089	-0,116234	0,9082
R-squared	0,218221	Mean dependent var		259504,2
Adjusted R-squared	0,169360	S,D, dependent var		402293,3
S,E, of regression	366648,0	Akaike info criterion		28,54401
Sum squared resid	4,30E+12	Schwarz criterion		28,67732
Log likelihood	-496,5201	Hannan-Quinn criter,		28,59003
F-statistic	4,466142	Durbin-Watson stat		2,228384
Prob(F-statistic)	0,019469			

Source : réalisé par l'auteur en utilisant Eviews 10.

Les valeurs p (Prob, F, Prob, Chi-Square) associées aux différents tests sont faibles (toutes inférieures à 0,05), ce qui signifie que nous rejetons l'hypothèse nulle d'homoscédasticité. Autrement dit, les erreurs résiduelles semblent avoir une différence entre les variances et il y a hétéroscédasticité dans ce modèle. En conséquence, des corrections s'imposent comme l'application d'une transformation logarithmique [log(SALES)] ou l'utilisation des erreurs standards robustes dans l'estimation de la régression.

- ✓ **Transformation logarithmique des variables :** Une autre solution consiste à transformer les variables en logarithmes, ce qui réduit l'effet des grandes valeurs et diminue l'hétéroscédasticité.

$$\log(\text{SALES}_t) = \beta_0 + \beta_1 \times \log(\text{ADVERTISING_EXPENSES}_t) + \beta_2 \times \log(\text{PRICE}_t) + \varepsilon_t$$

Tableau 15 : Régression des variables en logarithmes.

Dependent Variable: LOG(SALES)
Method: Least Squares
Date: 02/12/25 Time: 21:06
Sample: 2021M01 2023M11
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob,
C	2,468576	0,570668	4,325769	0,0001
LOG(ADVERTISING_EXPENSES)	0,798868	0,063900	12,50192	0,0000
LOG(PRICE)	-0,279483	0,084987	-3,288541	0,0025
R-squared	0,840758	Mean dependent var		7,686839
Adjusted R-squared	0,830806	S,D, dependent var		0,555919
S,E, of regression	0,228668	Akaike info criterion		-0,031277
Sum squared resid	1,673246	Schwarz criterion		0,102038
Log likelihood	3,547349	Hannan-Quinn criter,		0,014743
F-statistic	84,47609	Durbin-Watson stat		2,008392
Prob(F-statistic)	0,000000			

Source : réalisé par l'auteur en utilisant Eviews 10.

Les résultats après transformation montrent une absence de l'hétéroscédasticité (p-values du test de Breusch-Pagan après transformation > 0,10).

Tableau 16 : Résultat du test de Breusch-Pagan (après transformation des variables en logarithmes).

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0,211953	Prob, F(2,32)	0,8101
Obs*R-squared	0,457585	Prob, Chi-Square(2)	0,7955
Scaled explained SS	0,626694	Prob, Chi-Square(2)	0,7310

Test Equation:

Dependent Variable: RESID^2
 Method: Least Squares
 Date: 02/12/25 Time: 21:08
 Sample: 2021M01 2023M11
 Included observations: 35

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	-0,070195	0,224386	-0,312830	0,7564
LOG(ADVERTISING_EXPENSES)	0,006855	0,025125	0,272851	0,7867
LOG(PRICE)	0,019945	0,033417	0,596861	0,5548
R-squared	0,013074	Mean dependent var		0,047807
Adjusted R-squared	-0,048609	S,D, dependent var		0,087803
S,E, of regression	0,089912	Akaike info criterion		-1,898153
Sum squared resid	0,258694	Schwarz criterion		-1,764838
Log likelihood	36,21768	Hannan-Quinn criter,		-1,852133
F-statistic	0,211953	Durbin-Watson stat		2,091724
Prob(F-statistic)	0,810129			

Source : réalisé par l'auteur en utilisant Eviews 10.

- ✓ **Correction par les erreurs standards robustes (méthode de White) :** Cette méthode estime la même régression sous Eviews en choisissant dans les paramètres d'estimation et plus précisément dans le "Coefficient Covariance Matrix" l'option "Heteroskedasticity Consistent (White)" pour ajuster les erreurs standards des coefficients pour obtenir des erreurs standards robustes. Ainsi, les erreurs standards robustes permettent de garder les coefficients initiaux mais corrigent les p-values pour avoir des tests plus fiables.

Tableau 17 : Résultat du test de Breusch-Pagan (après correction par la méthode de White).

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	4,466142	Prob, F(2,32)	0,0195
Obs*R-squared	7,637735	Prob, Chi-Square(2)	0,0220
Scaled explained SS	7,452572	Prob, Chi-Square(2)	0,0241

Test Equation:

Dependent Variable: RESID^2
 Method: Least Squares
 Date: 02/12/25 Time: 21:15
 Sample: 2021M01 2023M11
 Included observations: 35

White-Hinkley (HC1) heteroskedasticity consistent standard errors and Covariance

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	-75853,44	145832,9	-0,520139	0,6065
ADVERTISING_EXPENSES	138,4544	51,71576	2,677218	0,0116
PRICE	-588,3865	4354,101	-0,135134	0,8934

R-squared	0,218221	Mean dependent var	259504,2
Adjusted R-squared	0,169360	S,D, dependent var	402293,3
S,E, of regression	366648,0	Akaike info criterion	28,54401
Sum squared resid	4,30E+12	Schwarz criterion	28,67732
Log likelihood	-496,5201	Hannan-Quinn criter,	28,59003
F-statistic	4,466142	Durbin-Watson stat	2,228384
Prob(F-statistic)	0,019469		

Source : réalisé par l'auteur en utilisant Eviews 10.

L'utilisation des erreurs standards robustes (White) pour corriger l'impact de l'hétéroscédasticité sans modifier le modèle initial ne permet pas de résoudre le problème de l'hétéroscédasticité (p-values du test de Breusch-Pagan après correction < 0,05).

Tableau 18 : Données sur les ventes, dépenses publicitaires et les prix de vente.

Date	Depenses_Publicitaires	Prix	Ventes
2021-01	1049,671	504,1773	469,9109
2021-02	1043,316	455,0923	644,9836
2021-03	1179,055	462,0077	265,4966
2021-04	1323,732	486,7944	711,3215
2021-05	1205,156	491,9122	564,602
2021-06	1262,301	474,8559	554,8713
2021-07	1500,778	463,4013	722,8425
2021-08	1476,743	453,9779	456,0434
2021-09	1410,195	424,7153	651,2105
2021-10	1568,542	434,1745	799,8814
2021-11	1525,087	433,6444	919,0406
2021-12	1581,998	458,2853	686,605
2022-01	1709,911	438,3009	713,9177
2022-02	1551,529	390,4535	694,6161
2022-03	1627,508	426,4817	906,0709
2022-04	1800,914	406,5841	924,522
2022-05	1813,003	395,033	813,4806
2022-06	2002,853	415,0907	1053,963
2022-07	1937,769	417,7628	957,8152
2022-08	1944,484	410,0542	1091,834
2022-09	2289,422	368,9299	1018,565
2022-10	2177,422	373,8158	1024,51
2022-11	2263,896	380,911	1053,685
2022-12	2171,811	388,0823	837,3811
2023-01	2316,99	353,2737	1201,624
2023-02	2439,664	353,4297	1258,7
2023-03	2370,615	329,3019	1187,379
2023-04	2580,427	321,7902	1251,68
2023-05	2539,936	356,2505	1012,577
2023-06	2627,973	361,4105	1229,259

2023-07	2654,115	327,1312	1266,302
2023-08	2956,656	342,9278	1316,733
2023-09	2827,222	324,3756	1385,434
2023-10	2779,943	298,5262	1476,778
2023-11	3025,112	312,9422	1898,279
2023-12	2877,916	330,7607	1473,579

Source : données fictives.

Conclusion

La variance de différence (hétéroscédasticité) est un problème fréquent dans les modèles de régression, particulièrement lorsque les données présentent des tendances ou des effets de taille. Les tests comme Breusch-Pagan, White et Goldfeld-Quandt permettent de la détecter. Des solutions comme les transformations, les moindres carrés pondérés et les erreurs standards robustes assurent une meilleure fiabilité des résultats.

Chapitre 6 : Tests aléatoires

Introduction

Les tests aléatoires sont essentiels dans les analyses économétriques pour vérifier la validité et la robustesse des modèles. Ils permettent notamment de s'assurer que les erreurs du modèle suivent une distribution normale, sont indépendantes, non corrélées et de variance constante. L'objectif de ces tests est de vérifier si les hypothèses classiques des modèles économétriques (comme le modèle des moindres carrés ordinaires – MCO) sont respectées.

Ce chapitre présente les principaux tests aléatoires, leurs méthodes de mise en œuvre et leurs interprétations pratiques. Certains de ces tests ont été déjà évoqués dans les chapitres précédents, mais leur importance dans la valide ou non des estimations économétriques par l'analyse des erreurs aléatoires. Elles doivent être :

- Normalement distribuées ;
- Non corrélées entre elles (pas d'autocorrélation) ;
- De variance constante (homoscédasticité).

Si ces conditions ne sont pas respectées, les estimateurs du modèle peuvent être biaisés, et les prédictions deviennent peu fiables.

1) Test de normalité des erreurs (Jarque-Bera)

1-1) Test et calcul de la statistique de Jarque-Bera

Le test de Jarque-Bera¹¹ vérifie si les erreurs résiduelles suivent une distribution normale. Ce test repose sur les coefficients d'asymétrie (Skewness) et d'aplatissement (Kurtosis). Il garantit la validité des inférences statistiques (tests de significativité, intervalles de confiance,...etc.).

Hypothèses :

- H_0 : Les erreurs suivent une distribution normale.
- H_1 : Les erreurs ne suivent pas une distribution normale.

Interprétation :

- Si la probabilité associée (p-value) est inférieure à 5 %, on rejette H_0 , indiquant une absence de normalité.
- Sinon, les erreurs peuvent être considérées comme normalement distribuées.

¹¹ Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55(2), 163-172. <https://doi.org/10.2307/1403192>

En effet, le test de Jarque-Bera utilise deux mesures de distribution des erreurs Skewness pour vérifier l'asymétrie autour de sa moyenne et Kurtosis pour vérifier l'aplatissement c'est-à-dire si la distribution est plus ou moins pointue qu'une distribution normale. La statistique de ce test, notée JB, est donnée par :

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

Avec :

n : Nombre d'observations

S : Coefficient d'asymétrie (Skewness)

K : Coefficient d'aplatissement (Kurtosis)

La statistique JB suit une loi Chi-deux (χ^2) avec 2 degrés de liberté.

1-2) Correction en cas de non-normalité

Si la normalité est rejetée, des corrections comme la transformation des variables ou l'élimination des valeurs extrêmes peuvent être envisagées :

- **Transformation des variables** : Appliquer une transformation logarithmique ou une différence première sur la variable dépendante. Si on considère la variable Y comme une variable indépendante, deux solutions sont possibles :

$$\checkmark Y' = \log(Y)$$

$$\checkmark Y' = Y_t - Y_{t-1}$$

- **Suppression des valeurs extrêmes** : Analyser les erreurs (résidus) standardisées et retirer les observations influentes.
- **Utilisation d'une estimation robuste** : Utiliser la méthode des moindres carrés généralisés (GLS) pour corriger l'hétéroscédasticité.

2) Test d'autocorrélation (Durbin-Watson)

2-1) Test et interprétation de la statistique de Durbin-Watson

Le test de Durbin-Watson¹², déjà évoqué dans le chapitre consacré à la régression linéaire et détaillé dans le chapitre traitant la corrélation des erreurs, mesure la corrélation entre les erreurs résiduelles consécutives. Une autocorrélation peut fausser les résultats et invalider les inférences statistiques, elle se produit lorsque les erreurs d'une période sont corrélées avec celles des périodes précédentes.

¹² Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression: I. *Biometrika*, 37(3-4), 409-428. <https://doi.org/10.2307/2332391>

Hypothèses :

- H_0 : Absence d'autocorrélation des erreurs.
- H_1 : Présence d'autocorrélation des erreurs.

Interprétation :

- Si la statistique de Durbin-Watson est proche de 2, il n'y a pas d'autocorrélation.
- Une valeur inférieure à 1,5 ou supérieure à 2,5 indique une possible autocorrélation.

2-2) Correction de l'autocorrélation

Si l'autocorrélation est détectée et confirmée par le test de Durbin-Watson, il faut envisager des corrections comme l'ajout de variables retardées ou l'utilisation de modèles autorégressifs:

- **Ajout d'une variable retardée** : Ajouter une variable Y_{t-1} (retard) dans le modèle pour capturer la dépendance temporelle.
- **Utilisation d'un modèle AR(1) (Autorégressif)** : Ajouter un terme autorégressif
$$e_t = \rho e_{t-1} + u_t$$
- **Régression avec Moindres Carrés Généralisés (GLS)** : Utiliser une estimation Newey-West pour corriger la variance des erreurs.
- **Différenciation des données** : Transformer les variables en différences premières
$$Y'_t = Y_t - Y_{t-1}$$

3) Test d'hétéroscédasticité (Breusch-Pagan, White)

3-1) Test et interprétation des statistiques de Breusch-Pagan et White

Ces tests, comme détaillés dans le chapitre consacré à la variance de différence, vérifient si la variance des erreurs est constante. L'hétéroscédasticité se produit lorsque la variance des erreurs d'un modèle de régression n'est pas constante. Cela viole l'hypothèse de base des Moindres Carrés Ordinaires (MCO) et peut entraîner des estimations biaisées des écarts-types des coefficients, rendant les tests statistiques non fiables.

Breusch-Pagan Test¹³ : Il repose sur une régression des carrés des résidus sur les variables explicatives pour détecter l'hétéroscédasticité.

Test de White¹⁴ : Une version plus générale, qui permet de détecter des formes non linéaires d'hétéroscédasticité.

¹³ Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5), 1287-1294. <https://doi.org/10.2307/1911963>

¹⁴ White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817-838. <https://doi.org/10.2307/1912934>

3-2) Correction de l'hétéroscédasticité

Si l'un de ces deux tests détecte une hétéroscédasticité (p-valeur < 5%), il faut appliquer une correction :

- **Utiliser les erreurs robustes de White pour** ajuster les écarts-types des coefficients :

Menu **Estimate Equation** → **Options** → Sélectionner **Heteroskedasticity Consistent Covariance (White)**.

- **Transformer les variables** : Utiliser le logarithme des variables explicatives :

$$\log(Y_t) = \beta_0 + \beta_1 \log(X_{1t}) + \beta_2 \log(X_{2t}) + \varepsilon_t$$

- **Appliquer la régression des moindres carrés généralisés (GLS)** : Si l'hétéroscédasticité suit une forme spécifique, la **GLS** peut corriger le problème.

4) Test de multicolinéarité (Variance Inflation Factor - VIF)

4-1) Test et statistique VIF

Bien que le VIF^{15} ne soit pas un test aléatoire proprement dit, il est souvent utilisé en complément pour détecter les relations linéaires entre les variables explicatives. Une forte multicolinéarité peut réduire la précision des estimateurs. Ce test et sa statistique ont été suffisamment développés dans le chapitre consacré à multicolinéarité.

4-2) Correction de la multicolinéarité

Si des valeurs $VIF > 10$, la multicolinéarité est confirmée, les corrections suivantes sont envisageables :

- **Supprimer une des variables fortement corrélées** (si elle n'apporte pas d'information essentielle) ;
- **Créer une nouvelle variable combinée** (par exemple, une somme ou une moyenne des variables corrélées) ;
- **Utiliser l'analyse en composantes principales (ACP)** pour regrouper les variables corrélées en un seul facteur ;
- **Standardiser les variables** (diviser par leur écart-type) pour réduire l'impact de l'échelle des données ;

¹⁵ Gujarati, D. N., & Porter, D. C. (2009). Basic econometrics (5th ed.). McGraw-Hill.

- Utiliser la régression ridge (Régression à pénalisation L2) pour minimiser l'impact de la multicolinéarité.

Exemple d'application sous Eviews (Vérification de la validité d'un modèle de régression) :

Pour modéliser les ventes en fonction des dépenses publicitaires et du prix du produit sur une période de 36 mois (2021-2023), nous allons travailler sur la base de données du tableau 19.

Tableau 19 : Données sur les ventes, dépenses publicitaires et les prix de vente.

Mois	VENTES	DEPENSES_PUBLICITAIRES	PRIX
2021M01	2463,07	4174	32,96
2021M02	2397,97	4507	28,74
2021M03	961,31	1860	38,36
2021M04	1145,89	2294	24,18
2021M05	800,77	2130	28,76
2021M06	952,31	2095	30,99
2021M07	2221,16	4772	33,68
2021M08	2536,88	4092	43,56
2021M09	1481,89	2638	25,99
2021M10	1065,65	3169	35,43
2021M11	838,42	1466	37,77
2021M12	1103,3	2238	21,39
2022M01	622,33	1330	38,23
2022M02	1468,8	2482	25,12
2022M03	1974,69	3135	21,95
2022M04	2687,01	4444	48,47
2022M05	1768,31	4171	48,97
2022M06	1882,13	3919	44,25
2022M07	2592,7	4735	29,14
2022M08	757,48	1130	22,93
2022M09	1260,42	2685	40,53
2022M10	2166,98	4380	33,2
2022M11	749,33	1769	23,66
2022M12	1336,08	3391	34,86
2023M01	1544,74	2515	21,03
2023M02	2907,6	4485	47,28
2023M03	1964,72	3853	27,76
2023M04	2123,2	3433	39,88
2023M05	1256	2215	29,35
2023M06	902,21	1955	35,6
2023M07	1842,06	3324	36,4
2023M08	1512,63	2184	25,55
2023M09	765,01	1459	49,09
2023M10	2942,34	4385	43,25
2023M11	259	1021	48,18
2023M12	1973,23	3300	46,84

Source : données fictives.

La régression des variables sous Eviews fournit les résultats du tableau 20.

Tableau 20 : Régression des ventes sur les dépenses publicitaires et les prix de vente.

Dependent Variable: VENTES
 Method: Least Squares
 Date: 02/13/25 Time: 21:47
 Sample: 2021M01 2023M12
 Included observations: 36

Variable	Coefficient	Std. Error	t-Statistic	Prob,
C	-112,1906	207,7408	-0,540051	0,5928
DEPENSES_PUBLICITAIRES	0,578574	0,042547	13,59849	0,0000
PRIX	-0,578357	5,482464	-0,105492	0,9166

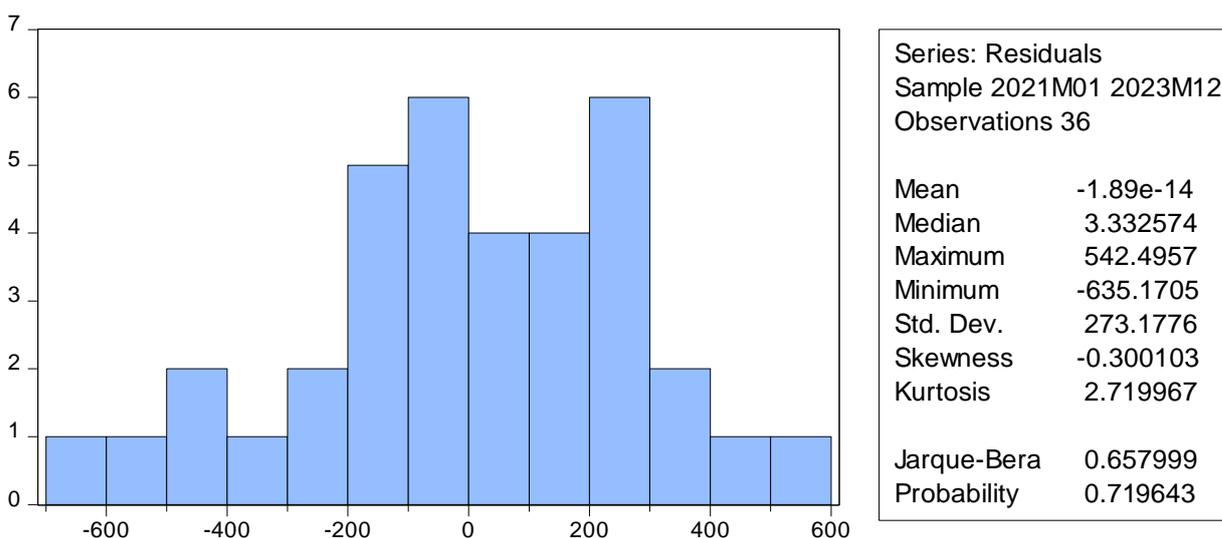
R-squared	0,854862	Mean dependent var	1589,656
Adjusted R-squared	0,846066	S,D, dependent var	717,0591
S,E, of regression	281,3340	Akaike info criterion	14,19662
Sum squared resid	2611911,	Schwarz criterion	14,32858
Log likelihood	-252,5391	Hannan-Quinn criter,	14,24267
F-statistic	97,18513	Durbin-Watson stat	2,037990
Prob(F-statistic)	0,000000		

Source : Elaboré par l'auteur à partir du tableau 19 en utilisant Eviews 10.

Les tests aléatoires pour valider son modèle :

- 1. Test de normalité des erreurs (Jarque-Bera) :** Les résultats du test de normalité sont présentés dans la figure 5.

Figure 5 : Résultats du test de normalité des erreurs de Jarque-Bera.



Source : réalisée par l'auteur en utilisant Eviews 10.

La probabilité associée (p-value) est supérieure à 5 %, on accepte H_0 selon laquelle les erreurs sont considérées comme normalement distribuées.

2. **Test d'autocorrélation (Durbin-Watson)** : La statistique de Durbin-Watson de 2,03 (tableau 20) est proche de 2, il n'y a pas donc d'autocorrélation.
3. **Test d'hétéroscédasticité (Breusch-Pagan)** : Les résultats du test de Breusch-Pagan sont présentés dans le tableau 21.

Tableau 21 : Résultats du test d'hétéroscédasticité de Breusch-Pagan.

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	3,607183	Prob, F(2,33)	0,0383
Obs*R-squared	6,458319	Prob, Chi-Square(2)	0,0396
Scaled explained SS	4,666943	Prob, Chi-Square(2)	0,0970

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 02/13/25 Time: 22:00

Sample: 2021M01 2023M12

Included observations: 36

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	-83416,21	66477,71	-1,254800	0,2184
DEPENSES_PUBLICITAIRES	28,27410	13,61517	2,076663	0,0457
PRIX	2079,696	1754,406	1,185413	0,2443
R-squared	0,179398	Mean dependent var		72553,07
Adjusted R-squared	0,129664	S,D, dependent var		96501,28
S,E, of regression	90027,77	Akaike info criterion		25,73328
Sum squared resid	2,67E+11	Schwarz criterion		25,86524
Log likelihood	-460,1990	Hannan-Quinn criter,		25,77934
F-statistic	3,607183	Durbin-Watson stat		2,315872
Prob(F-statistic)	0,038299			

Source : Réalisé par l'auteur en utilisant Eviews 10.

Les p-value sont faibles, on rejette H_0 . Il y a donc présence d'hétéroscédasticité qui peut-être corrigée par l'utilisation de l'estimation par moindres carrés généralisés (GLS).

Conclusion

Les tests aléatoires sont nécessaires pour valider et confirmer la robustesse des modèles économétriques. L'application de ces tests permet de garantir que les hypothèses fondamentales sont respectées et que les inférences réalisées sur les paramètres du modèle sont fiables. Eviews facilite l'exécution et l'interprétation de ces tests, rendant l'analyse plus rigoureuse et précise.

Chapitre 7 : Détection de la saisonnalité et méthodes de désaisonnalisation dans Eviews

Introduction

La saisonnalité est une composante essentielle de nombreuses séries temporelles économiques et financières. Elle se manifeste par des fluctuations récurrentes à intervalles réguliers, souvent causées par des événements périodiques comme les saisons climatiques, les habitudes de consommation ou les cycles d'activité économique.

L'objectif de la détection de la saisonnalité est d'identifier et de mesurer cette composante pour ensuite appliquer des techniques de désaisonnalisation, permettant ainsi une meilleure analyse des séries temporelles. Le logiciel Eviews¹⁶ offre plusieurs outils puissants pour détecter et supprimer la saisonnalité.

1) Détection de la saisonnalité

1-1) Analyse Graphique

Eviews permet de visualiser et de représenter les données sous différents formats pour détecter la saisonnalité :

- Graphiques de séries temporelles : Pour identifier les motifs récurrents visuellement.
- Histogrammes par période : Pour observer la distribution des valeurs sur différentes saisons.
- Correlogrammes : Disponible via la commande View → Correlogram, il permet d'analyser les coefficients d'autocorrélation à différents retards (lags).

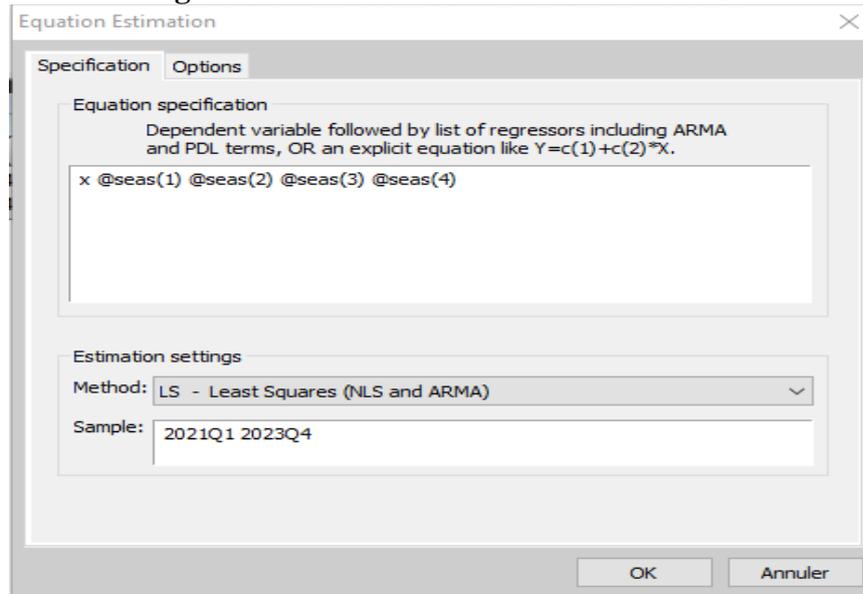
1-2) Test de saisonnalité dans Eviews

Le logiciel Eviews permet d'effectuer le test de saisonnalité, il suffit de saisir l'équation suivante : **Nom de la série @seas(1) @seas(2) @seas(3) @seas(4)** en utilisant la commande habituelle pour estimer une équation : Quick → estimate equation → écrire l'équation → cliquer sur OK (voir figure 6).

N/ B : Le nombre de coefficients @seas (...) est fonction de la périodicité de la série. Ainsi, pour une série semestrielle, nous saisissons deux coefficients, pour une série trimestrielle, nous saisissons quatre coefficients.

¹⁶ IHS Global Inc. (2021). Eviews 12 User's Guide. IHS Markit.

Figure 6 : Test de saisonnalité dans Eviews.



Source : réalisée par l’auteur en utilisant Eviews 10.

Pour valider la présence d’une saisonnalité, il faut que tous les coefficients du tableau 22 soient significatifs. C’est-à-dire que la valeur du t de Student calculée soit supérieure à la valeur tabulée, ou que la probabilité associée soit inférieure au seuil de risque de 5 %. Tous les coefficients ayant une probabilité égale à zéro ($p = 0$), ce qui est inférieur au risque de 0,05 (5 %), nous pouvons conclure que la série est saisonnière.

Tableau 22 : Résultats du test de saisonnalité.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@SEAS(1)	6.666667	0.927961	7.184212	0.0001
@SEAS(2)	9.000000	0.927961	9.698686	0.0000
@SEAS(3)	8.000000	0.927961	8.621054	0.0000
@SEAS(4)	26.00000	0.927961	28.01843	0.0000
R-squared	0.973052	Mean dependent var		12.41667
Adjusted R-squared	0.962947	S.D. dependent var		8.349832
S.E. of regression	1.607275	Akaike info criterion		4.048159
Sum squared resid	20.66667	Schwarz criterion		4.209795
Log likelihood	-20.28896	Hannan-Quinn criter.		3.988316
Durbin-Watson stat	2.591398			

Source : Réalisé par l’auteur en utilisant Eviews 10.

2) *Décomposition des séries temporelles dans Eviews*

Le logiciel Eviews permet de décomposer une série temporelle en différentes composantes afin de mieux comprendre son comportement. Cette décomposition vise à isoler les effets de tendance, de saisonnalité et des fluctuations aléatoires :

- **La tendance (Trend) T_t** : La tendance représente l'évolution à long terme d'une série temporelle. Elle reflète le comportement global des données en indiquant si la série suit une orientation haussière, baissière ou stable sur une période prolongée.
- **La saisonnalité S_t** : La saisonnalité correspond aux variations récurrentes qui se manifestent à intervalles réguliers dans le temps. Ce phénomène périodique peut être lié à des cycles naturels (saisons climatiques), économiques (ventes de fin d'année) ou institutionnels (échancier fiscal).
- **La composante résiduelle (bruit, erreurs)** : Les erreurs représentent les fluctuations accidentelles et imprévisibles de la série. Elles sont généralement dues à des événements ponctuels, tels que des crises économiques, des grèves ou des catastrophes naturelles, et sont considérées comme aléatoires.
- **La composante cyclique C_t** : La composante cyclique reflète les fluctuations économiques de moyen et long terme. Contrairement à la saisonnalité, ces cycles ne sont pas nécessairement réguliers et traduisent des phases successives de croissance et de récession qui influencent l'économie globale.

3) *Modèles de décomposition des séries temporelles*

3-1) *Modèle additif ou modèle multiplicatif*

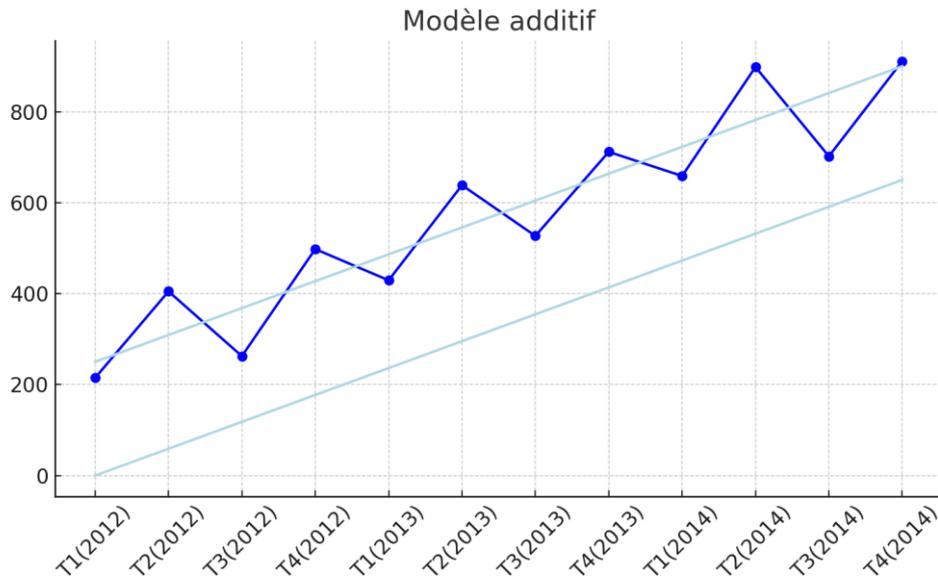
L'approche usuelle consiste à utiliser la méthode de décomposition classique, qui se base sur le modèle additif ou multiplicatif selon la nature des données :

- **Modèle additif** : Utilisé lorsque l'amplitude des variations saisonnières est relativement constante. Ce modèle suppose l'interdépendance des composantes précédemment citées :

$$X_t = T_t + S_t + C_t + e_t$$

Le modèle additif est engendré par deux lignes parallèles comme indiqué dans la représentation graphique de la figure 7.

Figure 7 : Représentation graphique du modèle additif.



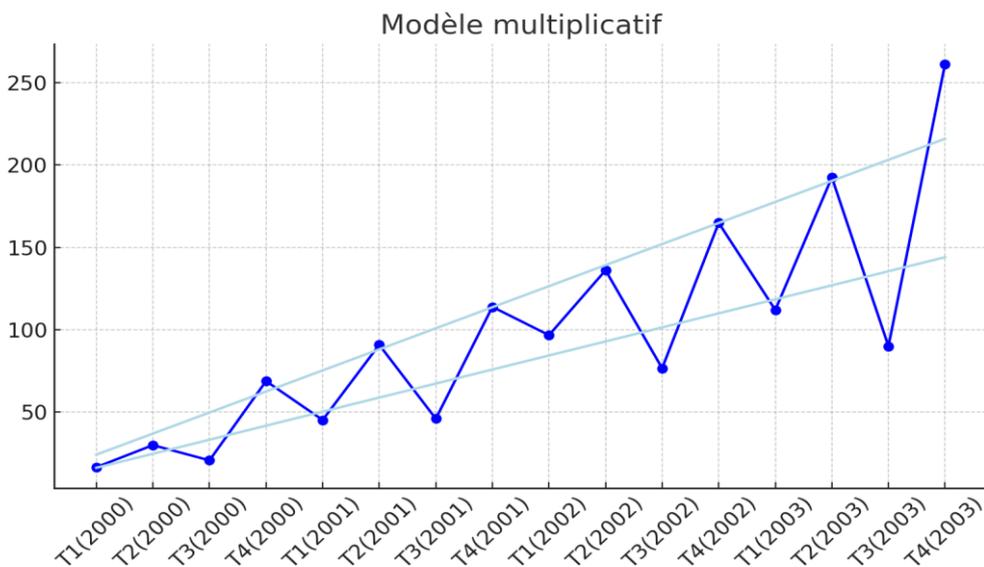
Source : réalisée par l’auteur en utilisant Eviews 10.

- **Modèle multiplicatif** : Employé lorsque l'effet saisonnier varie proportionnellement au niveau de la tendance :

$$X_t = T_t \times S_t \times C_t \times e_t$$

Le modèle multiplicatif est représenté graphiquement dans la figure 8.

Figure 8 : Représentation graphique du modèle multiplicatif.



Source : réalisée par l’auteur en utilisant Eviews 10.

3-2) Test de détermination du modèle (Buys-Ballot)

Le test de Buys-Ballot est une méthode statistique utilisée pour déterminer la nature du modèle d'une série temporelle (additif ou multiplicatif). Il repose sur l'analyse des moyennes et écarts types des saisons pour identifier la structure de la série. Ce test se base sur le calcul des moyennes et des écarts types par année. Le modèle est additif lorsque les moyennes et les écarts types sont indépendants. Dans le cas contraire, le modèle est multiplicatif. Pour effectuer ce test, on estime par la méthode des moindres carrés ordinaires les paramètres α et β du modèle suivant :

$$\sigma_i = \alpha + \beta \bar{x}_i + \varepsilon_i \quad i = \overline{1, n}$$

Si le coefficient β est significativement différent de zéro donc le modèle est multiplicatif. Dans le cas contraire (le coefficient β est égal à 0) le modèle est additif. Le logiciel Eviews peut appliquer ce test en suivant les étapes suivantes :

1. Importer la série temporelle et la structurer en mode panel avec les saisons ;
2. Calculer les moyennes et écarts types pour chaque saison via les statistiques descriptives ;
3. Construire un graphique de dispersion des écarts types en fonction des moyennes saisonnières ;
4. Estimer un modèle de régression linéaire entre écart type et moyenne :
 - Une pente significative indique un modèle multiplicatif.
 - Une pente non significative suggère un modèle additif.

4) Le test Dickey-Fuller Augmenté (ADF) et la modélisation VAR sous Eviews

4-1) Le test ADF

Les séries présentant une saisonnalité non corrigée sont souvent non stationnaires, car leurs moyennes et variances changent dans le temps. Si une série est non stationnaire en niveau mais devient stationnaire après différenciation saisonnière, cela indique la présence d'un effet saisonnier. Le test ADF est un test statistique utilisé pour déterminer si une série temporelle est stationnaire ou non. La stationnarité est une propriété essentielle pour éviter les fausses régressions¹⁷ dans l'analyse des données temporelles. Les hypothèses de ce test sont :

- H_0 La série temporelle contient une racine unitaire (elle est non stationnaire).
- H_1 : La série est stationnaire.

¹⁷ Enders, W. (2014). Applied Econometric Time Series (4th ed.). Wiley.

Le logiciel Eviews permet d'appliquer ce test en suivant les étapes :

1. **Importer les données** : Charger la série temporelle dans Eviews.
2. **Accéder au test ADF** :
 - Sélectionner la variable à tester.
 - Aller dans **View > Unit Root Test**.
 - Choisir **Augmented Dickey-Fuller**.
3. **Choisir la spécification** :
 - Avec constante seulement.
 - Avec tendance et constante.
 - Sans constante ni tendance.
4. **Analyser les résultats** :
 - Si la p-valeur est inférieure à 5 %, on rejette H_0 et on conclut que la série est stationnaire.
 - Si la p-valeur est supérieure à 5 %, on ne rejette pas H_0 et la série est considérée comme non stationnaire.
5. **Différenciation si besoin** : Si la série n'est pas stationnaire en niveau, tester sa première différence.

4-2) La modélisation VAR sous Eviews

Un VAR (Vector Autoregression) est une méthode utilisée pour modéliser l'interdépendance entre plusieurs séries temporelles. Si les variables incluent des composantes saisonnières, celles-ci doivent être désaisonnalisées ou différenciées pour éviter des estimations biaisées. Par ailleurs, une mauvaise prise en compte de la saisonnalité dans un modèle VAR peut fausser les relations dynamiques entre variables¹⁸. Le modèle VAR est un modèle économétrique multivarié permettant d'analyser les relations dynamiques entre plusieurs variables temporelles. Il suppose que chaque variable dépend de ses propres valeurs passées ainsi que de celles des autres variables du modèle. La modélisation VAR suppose :

- Les séries doivent être stationnaires (test ADF préalable) ;
- Le nombre de retards optimaux doit être déterminé (critères d'information AIC, BIC, HQ).

Le logiciel Eviews permet cette modélisation en suivant les étapes :

¹⁸ Enders, W. (2014). Op. cit.

1. **Importer les données** : Charger les variables dans Eviews.
2. **Créer un modèle VAR** :
 - Aller dans **Quick > Estimate VAR**.
 - Sélectionner les variables à inclure dans le modèle.
 - Déterminer le nombre de retards optimaux (via View → Lag Structure → Lag Length Criteria).
 - Choisir la méthode d'estimation et exécuter.
3. **Analyse des résultats** :
 - Examiner les coefficients et leurs significativités.
 - Analyser les tests de diagnostic (autocorrélation, normalité des résidus).
4. **Interprétation** :
 - Si un coefficient est significatif, il indique une relation dynamique entre les variables.
 - Évaluer l'impact des chocs à l'aide des fonctions de réponse impulsionnelle et de la décomposition de la variance.

5) Méthodes de désaisonnalisation dans Eviews

5-1) Méthode des Moyennes Mobiles

La méthode des moyennes mobiles est une technique simple et efficace pour éliminer la saisonnalité d'une série temporelle. Elle consiste à lisser les fluctuations saisonnières en prenant la moyenne de plusieurs observations consécutives. L'idée est de calculer une moyenne mobile centrée ou non centrée sur une fenêtre correspondant à la fréquence saisonnière.

- Si la saisonnalité est trimestrielle, on prend une moyenne mobile de 4 périodes.
- Si la saisonnalité est mensuelle, on utilise une moyenne mobile de 12 périodes.

Cela permet de neutraliser les effets saisonniers et d'extraire la tendance-cycle de la série.

L'application de la moyenne mobile sur une période définie pour désaisonnaliser une série temporelle suit les étapes suivantes :

1. Calcul de la moyenne mobile

- Choisir la fenêtre k correspondant à la saisonnalité (ex: 4 pour des données trimestrielles, 12 pour des données mensuelles),
- Appliquer la formule :

$$MM_t = \frac{X_{t-1} + X_t + X_{t+1}}{k}$$

- Pour une fenêtre paire, une moyenne mobile centrée est souvent utilisée.

2. Extraction de la composante saisonnière

- Comparer la série originale à la tendance obtenue pour identifier l'effet saisonnier.
- Calculer un indice saisonnier moyen pour chaque période saisonnière.

3. Correction de la saisonnalité

- Retirer les indices saisonniers de la série originale pour obtenir la série désaisonnalisée.
- Deux approches :
 - **Modèle additif** : $X_t - T_t = E_t$ (série désaisonnalisée)
 - **Modèle multiplicatif** : $X_t/S_t = T_t \times E_t$ (série désaisonnalisée)

Dans Eviews, on peut utiliser la méthode des moyennes mobiles avec la commande suivante :

1. Créer une moyenne mobile : **series MM = @movav(serie_brute, 4)**

"4" correspond à une moyenne mobile trimestrielle.

2. Calculer la composante saisonnière :
series composante_saisonniere = serie_brute - MM

3. Désaisonnaliser la série :

- Modèle additif : **series serie_desaison = serie_brute - composante_saisonniere**
- Modèle multiplicatif : **series serie_desaison = serie_brute / composante_saisonniere**

En fin, la méthode des moyennes mobiles est une approche courante pour éliminer la saisonnalité avant d'analyser une série temporelle. Elle est particulièrement utile avant d'appliquer des modèles de prévision comme le VAR, l'ARIMA ou les modèles de régression.

5-2) Ajustement par régression avec variables indicatrices

On peut modéliser la saisonnalité en utilisant une régression avec variables factices dans Eviews¹⁹ :

ls y c D1 D2 D3 ... D11

¹⁹ IHS Global Inc. (2021). Eviews 12 User's Guide. IHS Markit.

Où D1 à D11 sont des variables indicatrices pour chaque mois (le 12e étant la base de référence).

5-3) Méthode X-12-ARIMA et X-13-ARIMA-SEATS

Les méthodes X-12-ARIMA et X-13-ARIMA-SEATS²⁰ sont des techniques avancées utilisées pour la désaisonnalisation des séries temporelles. Elles sont développées par le Bureau of the Census des États-Unis et sont largement utilisées par les instituts statistiques et les économistes. La méthode X-12-ARIMA, basée sur un filtrage des tendances et des variations saisonnières, améliore les prévisions car elle prolonge la série pour permettre de mieux estimer la saisonnalité. Cette méthode, améliorée par l'intégration du Signal Extraction in ARIMA Time Series (SEATS) pour donner naissance à une autre méthode appelée X-13-ARIMA-SEATS. Le SEATS est basé sur une décomposition stochastique, ce qui permet d'obtenir des composantes plus précises et une séparation plus fine des tendances, cycles et irrégularités. Les deux méthodes suivent généralement les étapes suivantes :

1. **Ajustement des valeurs aberrantes et des jours ouvrables** : Correction des effets liés aux jours fériés, aux tendances inhabituelles.
2. **Modélisation ARIMA pour la prévision des données manquantes** : Ajustement d'un modèle ARIMA pour estimer et prolonger la série.
3. **Décomposition en composantes (Tendance, Saisonnière, Irrégulière)**
 - Extraction de la tendance-cycle.
 - Identification des effets saisonniers.
 - Élimination des fluctuations irrégulières.
4. **Désaisonnalisation finale** : Suppression de la composante saisonnière pour obtenir la série corrigée.

Le logiciel Eviews intègre X-12-ARIMA et X-13-ARIMA-SEATS directement dans ses options de désaisonnalisation, Il suffit de suivre les étapes suivantes :

1. **Importer la série temporelle** : Charger les données dans Eviews.
2. **Lancer la désaisonnalisation avec X-12-ARIMA** : en utilisant la Commande `seas ajuste_series = x12(series_brute)` et Eviews applique automatiquement X-12-ARIMA et génère une série désaisonnalisée.

²⁰ Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., & Chen, B. C. (1998). New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program. *Journal of Business & Economic Statistics*, 16(2), 127-177.

3. **Lancer la désaisonnalisation** avec X-13-ARIMA-SEATS en utilisant le menu **Proc** → **Seasonal Adjustment** → **X-13-ARIMA-SEATS** ou en syntaxe Eviews **seas ajuste_series = x13(series_brute)**

X-13 produit des résultats améliorés en combinant X-12 et SEATS.

4. **Interprétation des résultats :**

- Vérifier les diagnostics pour s'assurer de la bonne qualité de l'ajustement.
- Visualiser la série désaisonnalisée et la comparer avec la série brute.

5-4) Méthode STL (Seasonal-Trend Decomposition using Loess)

Bien que STL ne soit pas directement disponible dans Eviews, il est possible d'utiliser des scripts pour approcher cette méthode via des estimations locales.

5-5) Filtrage de Hodrick-Prescott (HP Filter)

Le filtre de Hodrick-Prescott²¹ est une méthode utilisée pour extraire la tendance d'une série temporelle en séparant la composante de long terme (tendance) des fluctuations cycliques et irrégulières. Il est particulièrement utilisé en macroéconomie pour analyser le cycle économique. En conséquence, cette méthode n'est pas développée dans ce cours.

Exemple d'application : Une série temporelle trimestrielle (T1 2000 à T4 2015) avec une tendance, une saisonnalité et une composante aléatoire. Cette série, notée serie_brute sera utilisée dans Eviews pour appliquer le test de Buys-Ballot et les méthodes de désaisonnalisation.

Tableau 23 : Données trimestrielle de la série brute.

	T1	T2	T3	T4	Moyenne	Ecart-type
2000	86,00921	103,0553	48,49698	142,2038	94,9413	38,8863
2001	155,5291	197,7304	38,77445	182,5299	143,641	72,0563
2002	206,9985	219,1425	105,425	182,6027	178,542	51,058
2003	180,4399	250,6118	115,9445	239,0731	196,517	61,88072
2004	213,0998	265,9509	163,1054	291,8122	233,492	57,22744

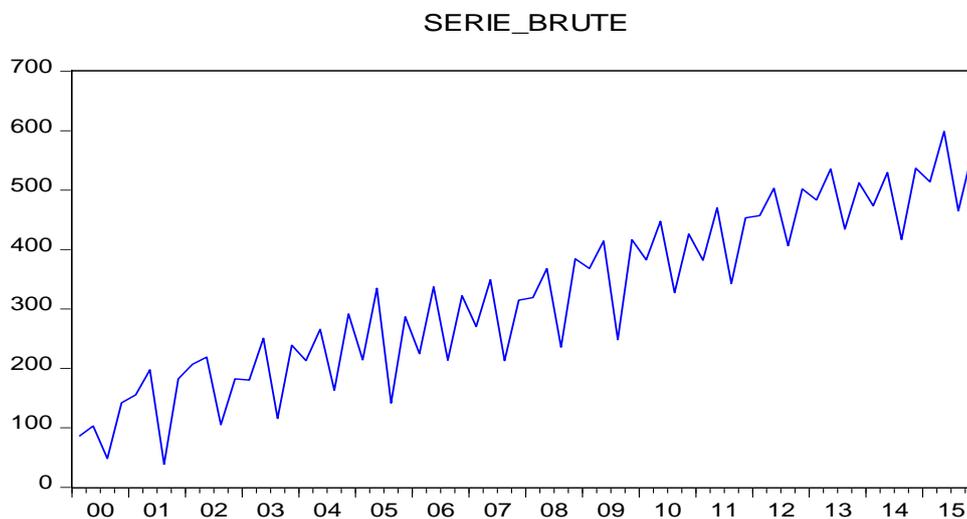
²¹ Hamilton, J. D. (2018). Why You Should Never Use the Hodrick-Prescott Filter. The Review of Economics and Statistics, 100(5), 831-843.

2005	214,6428	334,8228	141,893	286,8569	244,554	84,4037
2006	224,7665	337,335	213,9448	322,3641	274,603	64,23789
2007	270,3443	349,2507	213,1121	314,8711	286,895	58,84652
2008	319,3879	367,8621	235,8677	384,3109	326,857	66,62634
2009	368,2104	414,28	248,3472	416,6675	361,876	78,90332
2010	382,6847	447,3711	327,6121	426,1472	395,954	52,92144
2011	382,1935	470,3325	342,637	453,2815	412,111	60,0179
2012	457,1055	503,1762	406,5147	501,9732	467,192	45,78234
2013	483,2444	535,396	434,5498	512,393	491,396	43,49258
2014	473,7994	529,6192	416,9982	536,7336	489,288	55,80741
2015	514,1374	598,7792	465,0441	564,3486	535,577	58,47184

Source : Données fictives.

a/ Le graphe de la série brute, La série est-elle affectée par un mouvement saisonnier?

Figure 9 : Représentation graphique de la série brute.



Source : réalisée par l'auteur en utilisant Eviews 10.

Le graphe de la figure 9 montre la présence d'une tendance haussière et régulière au fil du temps (évolution positive de la variable étudiée). Il indique aussi des fluctuations récurrentes à intervalles réguliers, ce qui est typique d'un effet saisonnier. La série présente des pics et creux périodiques qui se répètent chaque année, probablement dus à une

influence trimestrielle. Il existe une certaine volatilité autour de la tendance : les valeurs fluctuent légèrement d'une période à l'autre. Cela indique la présence d'un bruit aléatoire qui peut être lié à des facteurs externes.

b / Le modèle de décomposition de la série brute :

D'après les résultats de la régression de l'écart-type sur la moyenne (Base_Buys_Ballot) du tableau 24, la probabilité associée au coefficient de la moyenne (0,57) est supérieure à 5%. La moyenne est indépendante de l'écart-type. Le modèle est donc additif.

Tableau 24 : Résultats de la régression de l'écart-type sur la moyenne (Base_Buys_Ballot).

Dependent Variable: ECART_TYPE
 Method: Least Squares
 Date: 02/19/25 Time: 19:56
 Sample: 2000 2015
 Included observations: 16

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	63,77941	8,252516	7,728482	0,0000
MOYENNE	-0,013607	0,023807	-0,571553	0,5767
R-squared	0,022802	Mean dependent var		59,41375
Adjusted R-squared	-0,046998	S,D, dependent var		12,21354
S,E, of regression	12,49725	Akaike info criterion		8,005363
Sum squared resid	2186,537	Schwarz criterion		8,101936
Log likelihood	-62,04290	Hannan-Quinn criter,		8,010308
F-statistic	0,326673	Durbin-Watson stat		1,899856
Prob(F-statistic)	0,576686			

Source : réalisé par l'auteur en utilisant Eviews 10.

c/ Test de saisonnalité

Quick → estimate equation → **serie_brute @seas(1) @seas(2) @seas(3) @seas(4)**
 → cliquer sur OK

Tableau 25 : Résultats du test de saisonnalité.

Dependent Variable: SERIE_BRUTE
 Method: Least Squares
 Date: 02/19/25 Time: 20:34
 Sample: 2000Q1 2015Q4
 Included observations: 64

Variable	Coefficient	Std, Error	t-Statistic	Prob,
@SEAS(1)	308,2871	34,15745	9,025471	0,0000
@SEAS(2)	370,2947	34,15745	10,84082	0,0000
@SEAS(3)	244,8917	34,15745	7,169496	0,0000

@SEAS(4)	359,8856	34,15745	10,53608	0,0000
R-squared	0,123856	Mean dependent var		320,8398
Adjusted R-squared	0,080048	S,D, dependent var		142,4502
S,E, of regression	136,6298	Akaike info criterion		12,73289
Sum squared resid	1120062,	Schwarz criterion		12,86782
Log likelihood	-403,4524	Hannan-Quinn criter,		12,78604
Durbin-Watson stat	0,055726			

Source : réalisé par l'auteur en utilisant Eviews 10.

Les résultats du tableau 25 confirment la présence d'une saisonnalité, la probabilité associée aux coefficients est égale à zéro ($p = 0$).

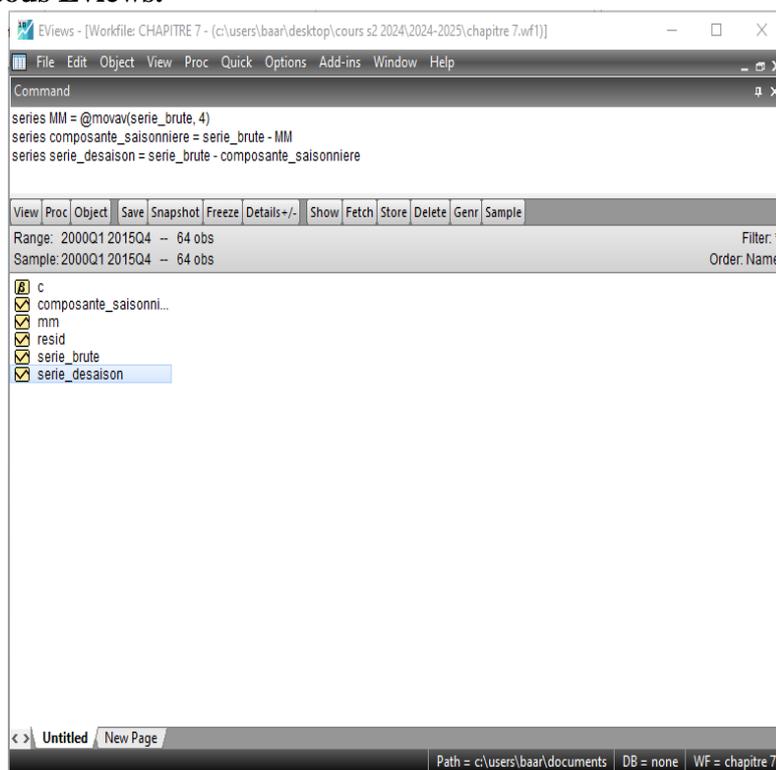
d/ **La dessaisonnalisation de la série brute (méthode des moyennes mobiles)** : Cette méthode lisse la série pour éliminer la saisonnalité.

Figure 10 : dessaisonnalisation de la série brute par la méthode des moyennes mobiles sous Eviews.

Créer une moyenne mobile centrée sur 4 périodes :
series MM = @movav(serie_brute, 4)

Calculer la composante saisonnière :
series composante_saisonniere = serie_brute - MM

Désaisonnaliser la série (Modèle additif) :
series serie_desaison = serie_brute - composante_saisonniere



Source : réalisée par l'auteur sous Eviews 10.

Enfin, cet exemple d'application dans Eviews nous a permis d'obtenir le tableau 26:

1. Une série désaisonnalisée (SERIE_DESAISON),
2. Une analyse des tendances et variations saisonnières,
3. Une modélisation via le test de Buys-Ballot et les moyennes mobiles.

Tableau 26 : Résultats de la désaisonnalisation.

	SERIE_BRUTE	MM	COMPOSANTE_SAISONNIE RE	SERIE_DESAISON
2000Q1	86,0092135835699			
2000Q2	103,055271282461			
2000Q3	48,4969785042555			
2000Q4	142,203837269106	94,9413251598481	47,2625121092579	94,9413251598481
2001Q1	155,52912632355	112,3213033448431	43,20782297870688	112,3213033448431
2001Q2	197,730447167899	135,9900973162026	61,74034985169638	135,9900973162026
2001Q3	38,7744510672046	133,5594654569399	-94,7850143897353	133,5594654569399
2001Q4	182,529928561569	143,6409882800557	38,88894028151333	143,6409882800557
2002Q1	206,998450200507	156,5083192492949	50,49013095121208	156,5083192492949
2002Q2	219,14248362596	161,8613283638102	57,28115526214986	161,8613283638102
2002Q3	105,42500897731	178,5239678413365	-73,09895886402651	178,5239678413365
2002Q4	182,602668771817	178,5421528938985	4,060515877918505	178,5421528938985
2003Q1	180,439911206542	171,9025181454073	8,537393061134736	171,9025181454073
2003Q2	250,611793737287	179,769845673239	70,84194806404801	179,769845673239
2003Q3	115,94450600901	182,399719931164	-66,45521392215399	182,399719931164
2003Q4	239,073054272862	196,5173163064253	42,55573796643677	196,5173163064253
2004Q1	213,099832022986	204,6822965105363	8,417535512449746	204,6822965105363
2004Q2	265,950853637597	208,5170614856138	57,43379215198325	208,5170614856138
2004Q3	163,105419731175	220,307289916155	-57,20187018498001	220,307289916155
2004Q4	291,812190197714	233,492073897368	58,320116300346	233,492073897368
2005Q1	214,642815593623	233,8778197900273	-19,23500419640428	233,8778197900273
2005Q2	334,822778869719	251,0958010980577	83,72697777166126	251,0958010980577
2005Q3	141,893012245147	245,7926992265508	-103,8996869814038	245,7926992265508
2005Q4	286,856894700344	244,5538753522083	42,30301934813576	244,5538753522083
2006Q1	224,766460847721	247,0847866657328	-22,31832581801174	247,0847866657328
2006Q2	337,335005845331	247,7128434096358	89,62216243569526	247,7128434096358
2006Q3	213,944824370977	265,7257964410933	-51,7809720701163	265,7257964410933
2006Q4	322,364144745406	274,6026089523588	47,76153579304725	274,6026089523588
2007Q1	270,344250488514	285,9970563625571	-15,65280587404305	285,9970563625571
2007Q2	349,250718957582	288,9759846406198	60,27473431696228	288,9759846406198
2007Q3	213,11207176592	288,7677964893556	-75,65572472343552	288,7677964893556
2007Q4	314,871138399841	286,8945449029643	27,97659349687678	286,8945449029643
2008Q1	319,387930963176	299,1554650216298	20,23246594154625	299,1554650216298
2008Q2	367,862059948348	303,8083002693213	64,05375967902676	303,8083002693213
2008Q3	235,867712157512	309,4972103672193	-73,62949820970729	309,4972103672193
2008Q4	384,31090594752	326,857152254139	57,45375369338098	326,857152254139
2009Q1	368,210442817924	339,062780217826	29,14766260009799	339,062780217826
2009Q2	414,279992494592	350,667263354387	63,61272914020504	350,667263354387
2009Q3	248,347239280034	353,7871451350175	-105,4399058549835	353,7871451350175
2009Q4	416,667501456873	361,8762940123558	54,79120744451728	361,8762940123558
2010Q1	382,68474481649	365,4948695119973	17,18987530449272	365,4948695119973
2010Q2	447,371054119146	373,7676349181358	73,60341920101024	373,7676349181358
2010Q3	327,612069359952	393,5838424381153	-65,97177307816326	393,5838424381153
2010Q4	426,147241281986	395,9537773943935	30,19346388759255	395,9537773943935
2011Q1	382,1934755542371	395,8309600788303	-13,63748452459322	395,8309600788303
2011Q2	470,332471559712	401,5713144389718	68,76115712074028	401,5713144389718
2011Q3	342,63695074615	405,3275347855213	-62,69058403937129	405,3275347855213
2011Q4	453,281544189034	412,1111105122833	41,17043367675075	412,1111105122833
2012Q1	457,105512053669	430,8391196371413	26,26639241652771	430,8391196371413
2012Q2	503,1762172332311	439,050056055521	64,12616117771006	439,050056055521
2012Q3	406,514666672981	455,0194850372288	-48,50481836424774	455,0194850372288
2012Q4	501,973201688322	467,1923994120508	34,78080227627123	467,1923994120508
2013Q1	483,244352816636	473,7271096027925	9,517243213843529	473,7271096027925
2013Q2	535,3960230017721	481,7820610449278	53,61396195684426	481,7820610449278
2013Q3	434,549783650168	488,7908402892245	-54,24105663905647	488,7908402892245

2013Q4	512,392986640253	491,3957865272073	20,99720011304567	491,3957865272073
2014Q1	473,799397241428	489,0345476334053	-15,2351503919773	489,0345476334053
2014Q2	529,619233226375	487,5903501895561	42,02888303681891	487,5903501895561
2014Q3	416,998228363391	483,2024613678618	-66,20423300447078	483,2024613678618
2014Q4	536,7336115604881	489,2876175979205	47,44599396256751	489,2876175979205
2015Q1	514,137381580824	499,3721136827695	14,7652678980545	499,3721136827695
2015Q2	598,779157516809	516,662094755378	82,11706276143104	516,662094755378
2015Q3	465,044128466508	528,6735697811573	-63,62944131464929	528,6735697811573
2015Q4	564,348600858029	535,5773171055426	28,77128375248651	535,5773171055426

Source : réalisé par l'auteur en utilisant Eviews 10.

Conclusion

L'analyse de la saisonnalité est une étape essentielle dans l'étude des séries temporelles, car elle permet de mieux comprendre les fluctuations périodiques et d'améliorer la qualité des prévisions. Dans ce chapitre, nous avons exploré différentes approches pour détecter et corriger la saisonnalité dans Eviews. Nous avons identifié la saisonnalité à travers des outils graphiques et les décompositions de séries, ainsi que par les tests de saisonnalité. Ces analyses nous ont permis de confirmer la présence de fluctuations saisonnières régulières dans certaines séries. Ensuite, nous avons étudié la désaisonnalisation des séries temporelles en nous basant sur leur décomposition (additive et multiplicative).

L'application de la désaisonnalisation dans Eviews a montré leur efficacité pour améliorer la lisibilité et l'analyse des données temporelles. Une série bien désaisonnalisée permet d'obtenir des modèles de prévision plus précis et d'éviter des biais liés à la saisonnalité dans les analyses prévisionnelles.

Enfin, après la détection et la correction de la saisonnalité, il est naturel de s'intéresser aux propriétés de dépendance temporelle des séries désaisonnalisées. C'est pourquoi le chapitre suivant abordera l'analyse des fonctions d'autocorrélation (ACF) et d'autocorrélation partielle (PACF), des outils fondamentaux pour identifier la structure des séries et sélectionner les modèles ARMA appropriés. Ainsi, la désaisonnalisation constitue une étape indispensable pour la modélisation des séries temporelles et l'application des techniques de prévision avancées que nous approfondirons dans les chapitres suivants.

Chapitre 8 : Fonction d'Autocorrélation (ACF) et Fonction d'Autocorrélation Partielle (PACF)

Introduction

En analyse des séries temporelles, l'étude de l'autocorrélation pour comprendre la relation entre les observations successives d'une même série est nécessaire. Il existe deux types d'autocorrélations :

- L'Autocorrélation (ACF) : mesure la corrélation entre une observation et ses valeurs passées à différents retards (lags).
- L'Autocorrélation Partielle (PACF) : mesure la corrélation entre une observation et une valeur passée en éliminant l'effet des autres retards intermédiaires.

Ces fonctions sont particulièrement utiles pour identifier l'ordre²² des processus AR(p), MA(q) et ARMA(p,q).

1) Définition et Formules

1-1) Fonction d'Autocorrélation (ACF)

L'autocorrélation à un retard k est définie comme :

$$\rho_k = \frac{Y_k}{Y_0}$$

Où :

- $Y_k = Cov(Y_t, Y_{t-k})$ est la covariance entre les observations à t et t-k.
- Y_0 est la variance de la série.

L'ACF est souvent représentée sous forme de graphique des autocorrélations pour détecter des schémas de dépendance temporelle.

1-2) Fonction d'Autocorrélation Partielle (PACF)

La PACF mesure l'autocorrélation entre Y_t et Y_{t-k} après avoir éliminé l'effet des retards intermédiaires 1,2,..., k-1. Elle est obtenue en estimant les coefficients de la régression :

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_k Y_{t-k} + \varepsilon_t$$

La PACF pour le retard k est alors le coefficient α_k

²² Brockwell, P. J., & Davis, R. A. (2016). Introduction to Time Series and Forecasting (3rd ed.). Springer.

2) Analyse de l'ACF et de la PACF pour les modèles AR, MA et ARMA

2-1) Modèle autorégressif (AR(p))

Un modèle AR(p) exprime une variable Y_t comme une combinaison linéaire de ses valeurs passées : $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$

Où :

ϕ_1 : Coefficients du modèle.

ε_t : Bruit blanc (erreur aléatoire indépendante et identiquement distribuée).

Le comportement des ACF et PACF pour un AR(p) :

- L'ACF décroît progressivement (exponentiellement ou sous forme oscillante).
- La PACF coupe brusquement après le lag p.

Exemple : Si la PACF coupe après $p = 2$, on peut suspecter un modèle AR(2).

2-2) Modèle moyenne mobile (MA(q))

Un modèle MA(q) exprime une variable Y_t comme une somme pondérée des erreurs passées :

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

Où θ_i représentent les coefficients du modèle.

Le comportement des ACF et PACF pour un MA(q) :

- L'ACF coupe brusquement après le lag q.
- La PACF décroît progressivement (exponentiellement ou oscillante).

Exemple : Si l'ACF coupe après $q = 3$, on peut suspecter un modèle MA(3).

2-3) Modèle ARMA(p,q)

Un modèle ARMA(p,q) combine les propriétés de AR(p) et MA(q) :

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

Le comportement des ACF et PACF pour un ARMA(p,q) : L'ACF et la PACF décroissent toutes les deux progressivement (exponentiellement ou oscillantes).

Exemple : Si ni l'ACF ni la PACF ne coupent brusquement, il faut ajuster un modèle ARMA.

Tableau 27 : Identification des modèles avec l'ACF et la PACF.

Modèle	ACF	PACF
AR(p)	Décroit progressivement	Coupe après p
MA(q)	Coupe après q	Décroit progressivement
ARMA(p,q)	Décroit progressivement	Décroit progressivement

Source : résumé par l'auteur à partir des définitions précédentes.

D'où la lecture du tableau 27 :

- Si la PACF coupe brusquement après un certain lag $p \rightarrow$ Modèle AR(p).
- Si l'ACF coupe brusquement après un certain lag $q \rightarrow$ Modèle MA(q).
- Si l'ACF et la PACF diminuent toutes les deux progressivement \rightarrow Modèle ARMA(p,q).

3) L'autocorrélation et les modèles AR(p), MA(q) et ARMA(p,q) dans Eviews

3-1) Générer et visualiser les correlogrammes

Dans Eviews, exécutez la commande : **show Y,correl** ou à partir du menu : View \rightarrow **Correlogram** \rightarrow **ACF et PACF**

3-2) Ajustement du modèle dans Eviews

L'ajustement des modèles dépend du type de l'autocorrélation détectée, dans Eviews²³ l'exécution de la commande selon le modèle :

- **Modèle AR(p)** : equation eq_ar,ls Y c ar(1) ar(2) ... ar(p)
 - ✓ equation eq_ar,ls : Définit une nouvelle équation appelée eq_ar.
 - ✓ Y : La variable dépendante (série temporelle).
 - ✓ c : L'intercept (constante du modèle).
 - ✓ ar(1) ar(2) ... ar(p) : Spécifie les termes auto-régressifs jusqu'à l'ordre p.
- **Modèle MA(q)** : equation eq_ma,ls Y c ma(1) ma(2) ... ma(q)
 - ✓ equation eq_ma,ls : Définit une nouvelle équation appelée eq_ma.
 - ✓ Y : La variable dépendante (série temporelle).
 - ✓ c : L'intercept (constante du modèle).
 - ✓ ma(1) ma(2) ... ma(q) : Spécifie les termes de moyenne mobile jusqu'à l'ordre q.
- **Modèle ARMA(p,q)** : equation eq_arma,ls Y c ar(1) ar(2) ma(1)
 - ✓ equation eq_arma,ls : Définit une nouvelle équation appelée eq_arma.

²³ IHS Global Inc. (2021). Eviews 12 User's Guide. IHS Markit.

- ✓ Y : La variable dépendante (série temporelle).
- ✓ c : L'intercept (constante du modèle).
- ✓ ar(1) ar(2) ..., ar(p) : Spécifie les termes auto-régressifs (AR) jusqu'à l'ordre p.
- ✓ ma(1) ma(2) ... ma(q) : Spécifie les termes de moyenne mobile (MA) jusqu'à l'ordre q.

Remarque : l'analyse et la validation des différents modèles suppose le respect des étapes suivantes :

- Vérifier la stationnarité avant d'utiliser l'ACF/PACF (test de Dickey-Fuller par exemple).
- Lire l'ACF et PACF pour identifier le modèle AR(p), MA(q) ou ARMA(p,q).
- Estimer le modèle sous Eviews et valider avec des critères statistiques (AIC, BIC).

Application dans Eviews :

L'analyse de l'ACF et de la PACF pour identifier un modèle autorégressif ainsi que son application sous Eviews en utilisant les données de l'exemple d'application du chapitre précédent. Etant donné que la série de données désaisonnalisée déjà existante dans l'espace de travail Eviews, nous allons suivre les étapes suivantes :

1. Vérification de la stationnarité de la série désaisonnalisée (testons la stationnarité avec **Dickey-Fuller Augmenté (ADF)**):

A partir du menu **Quick** → **Group Statistics** → **Unit Root Test**, nous avons obtenu le résultat du tableau 28.

Tableau 28 : Test de stationnarité (ADF) de la variable serie_desaison en niveau.

Null Hypothesis: SERIE_DESAISON has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 1 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob,*
Augmented Dickey-Fuller test statistic	-3,139345	0,1070
Test critical values:		
1% level	-4,121303	
5% level	-3,487845	
10% level	-3,172314	

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

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(SERIE_DESAISON)
 Method: Least Squares
 Date: 02/24/25 Time: 22:35

Sample (adjusted): 2001Q2 2015Q4
 Included observations: 59 after adjustments

Variable	Coefficient	Std, Error	t-Statistic	Prob,
SERIE_DESAISON(-1)	-0,271691	0,086544	-3,139345	0,0027
D(SERIE_DESAISON(-1))	0,227833	0,127734	1,783650	0,0800
C	28,10180	7,209161	3,898069	0,0003
@TREND("2000Q1")	1,904438	0,609253	3,125858	0,0028
R-squared	0,161401	Mean dependent var		7,173831
Adjusted R-squared	0,115659	S,D, dependent var		6,907128
S,E, of regression	6,495422	Akaike info criterion		6,645461
Sum squared resid	2320,478	Schwarz criterion		6,786311
Log likelihood	-192,0411	Hannan-Quinn criter,		6,700443
F-statistic	3,528520	Durbin-Watson stat		2,074166
Prob(F-statistic)	0,020651			

Source : réalisé par l'auteur en utilisant Eviews 10.

Interprétation : p-value > 0,05, la série doit être différenciée car une série qui n'est pas stationnaire doit être transformée (différenciée) avant d'appliquer un modèle ARMA.

Tableau 29 : Test de stationnarité (ADF) de la variable serie_desaison en première différence.

Null Hypothesis: D(SERIE_DESAISON) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob,*
Augmented Dickey-Fuller test statistic	-6,862710	0,0000
Test critical values:		
1% level	-4,121303	
5% level	-3,487845	
10% level	-3,172314	

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

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(SERIE_DESAISON,2)
 Method: Least Squares
 Date: 02/24/25 Time: 22:39
 Sample (adjusted): 2001Q2 2015Q4
 Included observations: 59 after adjustments

Variable	Coefficient	Std, Error	t-Statistic	Prob,
D(SERIE_DESAISON(-1))	-0,896715	0,130665	-6,862710	0,0000
C	6,476998	2,288735	2,829946	0,0065
@TREND("2000Q1")	-0,001837	0,053509	-0,034330	0,9727
R-squared	0,457343	Mean dependent var		-0,177563
Adjusted R-squared	0,437962	S,D, dependent var		9,324037
S,E, of regression	6,990154	Akaike info criterion		6,776391
Sum squared resid	2736,286	Schwarz criterion		6,882029
Log likelihood	-196,9035	Hannan-Quinn criter,		6,817628

F-statistic	23,59796	Durbin-Watson stat	2,028171
Prob(F-statistic)	0,000000		

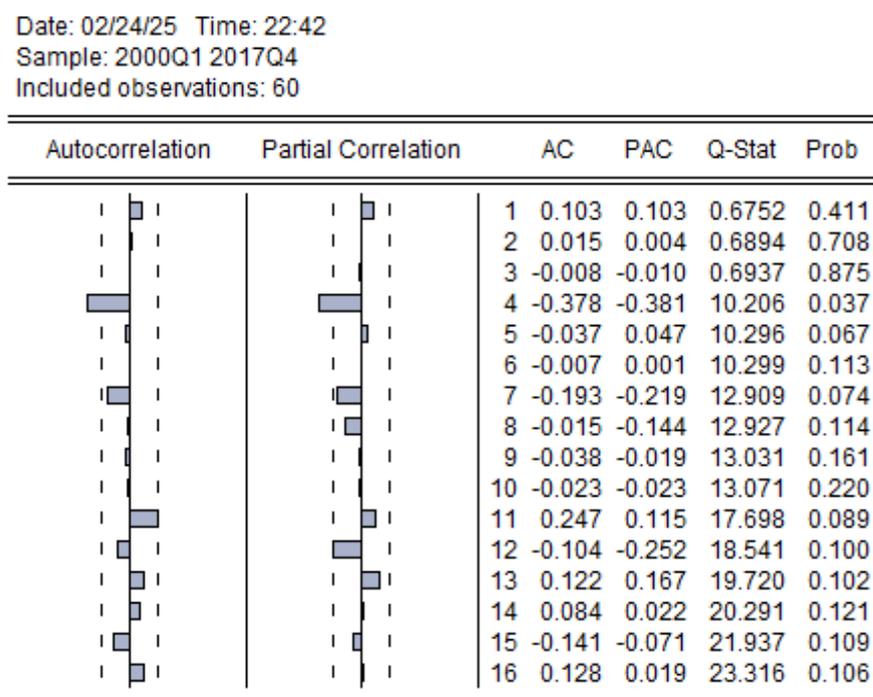
Source : réalisé par l'auteur en utilisant Eviews 10.

Interprétation : p-value < 0,05, la série est stationnaire en premières différences,

2. Analyse de l'ACF et de la PACF

Générer et visualiser les correlogrammes (à partir du menu: View → Correlogram → ACF et PACF), nous avons obtenu les résultats de la figure 11.

Figure 11 : Correlogramme des erreurs.



Source : réalisée par l'auteur en utilisant Eviews 10.

La figure 11 affiche les autocorrélations (ACF) et autocorrélations partielles (PACF) pour la série désaisonnalisée différenciée :

- L'ACF (première colonne) mesure la corrélation entre la série différenciée et ses valeurs décalées (lags). L'ACF montre une décroissance rapide après le lag 2, avec un pic négatif significatif au lag 4.

- La PACF (deuxième colonne) permet d'identifier les lags pertinents en éliminant l'effet des lags intermédiaires. La PACF présente une coupure nette après le lag 4, ce qui suggère une structure autorégressive.

Les probabilités associées (Prob) sont non significatives pour la plupart des retards, indiquant une absence d'autocorrélation. La PACF coupe après $p=4$, ce qui suggère l'estimation d'un modèle AR(4) et l'ACF ne présente pas une coupure nette mais oscille faiblement, ce qui rend l'identification d'un modèle MA seule difficile.

3. **Estimation du modèle approprié :** Comme suggéré par l'analyse du correlogramme précédent, nous allons estimer un AR(4) et vérifier la significativité des coefficients.

Tableau 30 : Estimation d'un modèle AR(4).

Dependent Variable: D(SERIE_DESAISON)
 Method: ARMA Maximum Likelihood (OPG - BHHH)
 Date: 02/24/25 Time: 23:33
 Sample: 2001Q1 2015Q4
 Included observations: 60
 Convergence achieved after 22 iterations
 Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	7,161039	0,659283	10,86187	0,0000
AR(1)	0,093718	0,104335	0,898245	0,3730
AR(2)	0,021722	0,131847	0,164751	0,8698
AR(3)	0,028472	0,134819	0,211189	0,8335
AR(4)	-0,445812	0,140180	-3,180273	0,0024
SIGMASQ	38,63948	10,23622	3,774779	0,0004
R-squared	0,192070	Mean dependent var		7,343933
Adjusted R-squared	0,117262	S,D, dependent var		6,973943
S,E, of regression	6,552309	Akaike info criterion		6,707119
Sum squared resid	2318,369	Schwarz criterion		6,916553
Log likelihood	-195,2136	Hannan-Quinn criter,		6,789040
F-statistic	2,567497	Durbin-Watson stat		1,914905
Prob(F-statistic)	0,037222			
Inverted AR Roots	,61-,56i	,61+,56i	-,56+,58i	-,56-,58i

Source : réalisé par l'auteur en utilisant Eviews 10.

Interprétation du modèle AR(4)

Lecture des coefficients et signification statistique

- **Constante (C = 7,161) :** La constante est significative ($p\text{-value} = 0,0000$). Cela signifie que la série a une tendance moyenne positive,
- **AR(1) = 0,0937, AR(2) = 0,0217 et AR(3) = 0,0285 :** Ces coefficients ne sont pas significatifs ($p\text{-values} > 0,05$), donc leurs effets sont statistiquement négligeables.

- **AR(4) = -0,4458** : Ce coefficient est statistiquement significatif (p-value = 0,0024). Cela indique que le quatrième retard de la série joue un rôle important dans la dynamique du processus.

Qualité du modèle

- **R² = 0,1920, Ajusté R² = 0,1172** : Ces valeurs sont relativement faibles, ce qui signifie que le modèle n'explique qu'environ 19 % de la variance de la série.
- **Akaike (6,7071), Schwarz (6,9166), Hannan-Quinn (6,7890)** : Ces critères peuvent servir à comparer ce modèle à d'autres modèles ARMA.
- **Durbin-Watson = 1,9149** : Cette valeur est proche de 2, ce qui suggère une absence d'autocorrélation significative des résidus.

En conclusion, le modèle AR(4) est approprié, mais sa capacité explicative reste faible. Le terme AR(4) est significatif, mais les autres termes AR ne le sont pas, Il est indispensable de vérifier la validation du modèle estimé.

4. Validation du Modèle

Le résultat, selon lequel les erreurs suivent un bruit blanc, obtenu par le test de Durbin-Watson (DW) est confirmé par l'analyse de l'autocorrélation des erreurs (figure 12).

Figure 12 : Correlogramme des erreurs.

Date: 02/24/25 Time: 23:40
Sample: 2000Q1 2017Q4
Included observations: 60
Q-statistic probabilities adjusted for 4 ARMA terms

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.024	0.024	0.0372	
		2	0.002	0.001	0.0374	
		3	-0.109	-0.109	0.8063	
		4	-0.028	-0.023	0.8585	
		5	0.022	0.025	0.8925	0.345
		6	0.038	0.026	0.9912	0.609
		7	-0.081	-0.090	1.4559	0.692
		8	-0.246	-0.243	5.7708	0.217
		9	0.030	0.049	5.8346	0.323
		10	-0.003	-0.016	5.8355	0.442
		11	0.188	0.141	8.5317	0.288
		12	-0.127	-0.156	9.7909	0.280
		13	0.088	0.118	10.404	0.319
		14	0.052	0.095	10.624	0.388
		15	-0.117	-0.204	11.765	0.382
		16	0.057	0.028	12.042	0.442

Source : réalisée par l'auteur en utilisant Eviews 10.

Interprétation du correlogramme des résidus :

Ce correlogramme permet de vérifier si le modèle AR(4) ajuste bien les données en analysant l'autocorrélation des résidus :

✓ **Absence d'Autocorrélation :**

- Toutes les valeurs d'autocorrélation (AC) et d'autocorrélation partielle (PAC) sont proches de zéro.
- Aucune valeur de l'ACF ou PACF ne dépasse les bandes de l'intervalle de confiance.
- Les p-values du test de Ljung-Box (Prob) sont toutes supérieures à 0,05, ce qui indique que les résidus sont non corrélés et suivent un bruit blanc.

En conclusion, le modèle AR(4) est bien spécifié, Il capture correctement la structure des données et les résidus étant du bruit blanc, aucune autre structure temporelle significative ne reste à modéliser.

Conclusion

L'analyse des fonctions d'autocorrélation (ACF) et d'autocorrélation partielle (PACF) est nécessaire pour comprendre la structure de dépendance dans une série temporelle. L'ACF mesure la corrélation entre les observations d'une série à différents retards (lags), permettant d'identifier la persistance des effets dans le temps. La PACF, quant à elle, permet d'isoler la relation directe entre une observation et un lag donné, en éliminant l'influence des lags intermédiaires.

Ces outils sont particulièrement utiles dans la méthodologie Box-Jenkins qui sera détaillée dans le dixième chapitre. En effet, la maîtrise de l'ACF et de la PACF permet d'affiner la modélisation des séries temporelles, d'optimiser les prévisions et d'améliorer l'interprétation des dynamiques temporelles. Une utilisation rigoureuse de ces outils contribue à la sélection d'un modèle statistique pertinent et efficace.

Chapitre 9 : Tests de Stabilité

Introduction

L'analyse de la stabilité des coefficients d'un modèle économétrique est une étape nécessaire pour garantir la robustesse et la fiabilité des estimations. Une instabilité structurelle peut fausser les prédictions et invalider les conclusions tirées du modèle²⁴. Ce chapitre présente les principaux tests de stabilité utilisés en économétrie et leur application sous Eviews²⁵.

1) Le Test de Chow

Le test de Chow est utilisé pour détecter une rupture structurelle à un moment précis dans une série temporelle.

1-1) Hypothèses du test

- Hypothèse nulle (H_0) : Il n'y a pas de rupture structurelle, les coefficients du modèle sont constants.
- Hypothèse alternative (H_1) : Il existe une rupture structurelle, les coefficients changent après un certain point.

1-2) Méthodologie

Le test consiste à diviser l'échantillon en deux sous-échantillons à partir d'un point de rupture supposé. Trois régressions sont effectuées :

1. Sur l'ensemble des données.
2. Sur le premier sous-échantillon.
3. Sur le second sous-échantillon.

La statistique de test de Chow est calculée comme suit :

$$F = \frac{(S_{total} - (S_1 + S_2))/k}{(S_1 + S_2)/(N - 2k)}$$

Avec :

- S_{total} la somme des carrés des résidus du modèle global.
- S_1 et S_2 les sommes des carrés des résidus des sous-échantillons.
- k le nombre de paramètres.
- N la taille totale de l'échantillon.

Si la statistique F dépasse une valeur critique, on rejette et on conclut qu'il y a une rupture structurelle.

²⁴ Greene, W. H. (2018). *Econometric Analysis* (8th ed.). Pearson.

²⁵ IHS Global Inc. (2021). *Eviews 12 User's Guide*. IHS Markit.

1-3) Application sous Eviews

1. Charger les données et exécuter une régression initiale via **Quick** → **Estimate Equation**.
2. Identifier le point de rupture potentiel et diviser l'échantillon en deux sous-échantillons.
3. Effectuer les régressions pour chaque sous-échantillon.
4. Aller dans **View** → **Stability Tests** → **Chow Breakpoint Test**.
5. Interpréter les résultats fournis par Eviews.

2) Tests de Stabilité des Résidus : CUSUM et CUSUMSQ

Ces tests sont basés sur l'analyse des résidus cumulés et permettent de détecter des instabilités progressives.

2-1) Test de CUSUM

Le test de CUSUM (Cumulative Sum) examine la somme cumulée des résidus normalisés. Si cette somme sort d'une bande de l'intervalle de confiance prédéfinie, une instabilité est détectée.

2-2) Test de CUSUMSQ

Le test CUSUMSQ (CUSUM of Squares) est une extension du CUSUM et est particulièrement utile pour détecter des changements graduels dans la variance des résidus.

2-3) Application sous Eviews

1. Estimer la régression sous **Quick** → **Estimate Equation**.
2. Aller dans **View** → **Stability Tests** et sélectionner **CUSUM Test** ou **CUSUM of Squares Test**.
3. Vérifier si la courbe reste dans la bande de l'intervalle de confiance à 95%.

3) Test de Breakpoint de Bai-Perron

Le test de Bai-Perron est une méthode avancée permettant de détecter plusieurs points de rupture structurelle.

3-1) Méthodologie

Il repose sur l'estimation de modèles à régimes multiples et identifie les moments où les coefficients changent significativement. La procédure utilise des tests séquentiels et une approche de rééchantillonnage (bootstrapping) pour estimer les points de rupture optimaux.

3-2) Application sous Eviews

1. Charger les données et exécuter une régression initiale.
2. Aller dans **View** → **Stability Tests** → **Multiple Breakpoint Test**.

3. Spécifier le nombre maximum de ruptures et exécuter le test.
4. Analyser les résultats et identifier les périodes de rupture.

Exemple d'application :

Dans cette application, nous allons effectuer les tests de stabilité sur des données financières (tableau 31), à savoir :

- Indice_Boursier (évolution d'un indice boursier avec fluctuations) ;
- Taux_Interet (taux d'intérêt mensuel entre 1% et 5%) ;
- Inflation (variation de l'inflation entre 1% et 3%) ;
- Taux_Change (évolution du taux de change avec variations aléatoires).

Tableau 31 : Données financières.

	INDICE_BOURSIER	TAUX_INTERET	INFLATION	TAUX_CHANGE
2000M01	1001,493	2,128138	2,143225	1,087461
2000M02	1001,717	1,709758	1,559958	1,12371
2000M03	1003,512	4,002459	2,538986	1,137865
2000M04	1007,058	4,227339	1,374087	1,126615
2000M05	1007,09	4,962021	1,647358	1,139263
2000M06	1007,122	2,650471	1,850873	1,158715
2000M07	1010,78	2,488072	2,015221	1,171151
2000M08	1012,815	4,105652	1,484819	1,139746
2000M09	1012,376	2,363214	1,229674	1,125204
2000M10	1013,961	4,723029	2,22124	1,120253
2000M11	1013,534	4,433651	1,577261	1,118764
2000M12	1013,103	2,715976	2,162476	1,131178
2001M01	1014,087	4,003484	1,308725	1,134732
2001M02	1010,76	4,018171	1,96228	1,108025
2001M03	1007,81	1,412495	2,065179	1,115629
2001M04	1007,186	4,610212	1,103647	1,127841
2001M05	1005,66	3,021009	1,673209	1,139037
2001M06	1006,789	4,30583	1,268829	1,160652
2001M07	1005,473	2,280198	1,12675	1,177331
2001M08	1003,148	4,582093	2,97992	1,186514
2001M09	1006,579	2,556807	1,644708	1,185111
2001M10	1006,628	1,043351	2,619749	1,151892
2001M11	1007,263	4,621528	1,509281	1,160484
2001M12	1004,913	1,365147	2,363005	1,164638
2002M01	1004,325	2,277255	2,520456	1,170069
2002M02	1005,046	4,800248	2,191277	1,144534
2002M03	1003,244	4,802429	1,943152	1,122913
2002M04	1004,496	3,293752	1,823682	1,143976
2002M05	1003,795	3,527349	1,697737	1,143185
2002M06	1003,711	2,793782	2,859058	1,156815

2002M07	1003,008	2,172843	2,661239	1,157382
2002M08	1007,212	2,314658	2,930054	1,157977
2002M09	1007,685	3,690074	1,248594	1,176742
2002M10	1006,07	4,009498	2,461735	1,166422
2002M11	1008,215	4,166316	2,876681	1,168344
2002M12	1006,273	4,158473	1,362466	1,159098
2003M01	1007,191	1,364824	1,132993	1,150408
2003M02	1003,772	2,977681	2,482241	1,144225
2003M03	1001,615	1,230235	2,148946	1,148668
2003M04	1002,509	3,198116	2,683658	1,139093
2003M05	1004,486	2,766122	1,279545	1,164208
2003M06	1005,329	4,550817	2,590535	1,146316
2003M07	1005,597	2,40366	1,403255	1,142578
2003M08	1005,495	1,468268	1,327312	1,133784
2003M09	1003,038	1,571967	1,328532	1,162723
2003M10	1002,098	4,046043	2,629149	1,166654
2003M11	1001,677	3,472872	2,330394	1,187291
2003M12	1004,291	1,404491	2,046131	1,15758
2004M01	1005,479	1,336427	1,717661	1,162921
2004M02	1002,453	3,803877	2,754401	1,180714
2004M03	1003,601	1,291052	1,78489	1,182359
2004M04	1003,331	4,28744	2,633199	1,203669
2004M05	1002,477	3,824969	1,87827	1,193323
2004M06	1004,2	1,325395	1,753889	1,22151
2004M07	1006,762	1,339351	1,92536	1,267488
2004M08	1009,125	4,946558	1,602756	1,260231
2004M09	1007,946	2,497083	2,495219	1,251321
2004M10	1007,828	2,482569	2,005441	1,280389
2004M11	1008,99	4,251198	1,464425	1,31198
2004M12	1011,441	4,788994	2,799149	1,301523
2005M01	1010,983	4,944004	1,767782	1,293119
2005M02	1011,112	4,013513	2,087106	1,287484
2005M03	1009,399	2,505038	2,812944	1,260595
2005M04	1007,507	1,334003	2,248476	1,242222
2005M05	1009,632	4,108588	1,233796	1,222139
2005M06	1012,844	3,233617	2,879664	1,206783
2005M07	1013,2	2,696888	2,255416	1,206089
2005M08	1015,707	4,625418	1,669811	1,210774
2005M09	1016,931	1,44479	1,278544	1,241784
2005M10	1016,14	2,9705	2,58805	1,221816
2005M11	1017,363	1,045415	2,240146	1,241503
2005M12	1020,939	2,874643	2,066922	1,237223
2006M01	1021,368	1,225213	2,787785	1,236234
2006M02	1024,997	1,475272	2,577194	1,24973
2006M03	1020,257	1,470105	1,30335	1,227276
2006M04	1022,401	3,596841	1,623444	1,234924

2006M05	1023,075	3,98418	1,496978	1,238253
2006M06	1022,977	3,333475	2,487893	1,248102
2006M07	1023,661	4,84869	1,067065	1,253885
2006M08	1020,186	2,499482	2,139779	1,302991
2006M09	1020,246	2,142848	2,524917	1,290237
2006M10	1021,46	4,474397	2,753531	1,279617
2006M11	1024,916	1,894383	1,684163	1,267154
2006M12	1024,38	4,85289	2,642515	1,256044
2007M01	1023,263	1,048618	1,221263	1,243297
2007M02	1022,759	4,879515	2,692905	1,267077
2007M03	1025,09	1,17264	1,254977	1,295487
2007M04	1026,248	4,564572	1,794575	1,284072
2007M05	1025,688	3,110804	2,594591	1,267425
2007M06	1027,215	4,971859	1,299835	1,276853
2007M07	1027,909	1,295186	1,458503	1,265809
2007M08	1030,346	3,215417	2,444505	1,278467
2007M09	1029,442	4,87721	2,440073	1,282526
2007M10	1029,287	3,092391	2,282295	1,252211
2007M11	1029,002	3,517595	2,387897	1,283161
2007M12	1026,575	3,782995	2,085449	1,319079
2008M01	1027,668	2,818164	1,503598	1,306823
2008M02	1028,69	3,510232	1,691392	1,299069
2008M03	1029,2	3,337257	1,363195	1,304786
2008M04	1029,231	4,604632	2,816901	1,311475
2008M05	1026,9	1,181786	2,166784	1,324646
2008M06	1026,559	2,123853	1,801703	1,36485
2008M07	1026,373	4,801646	1,924012	1,361311
2008M08	1025,269	4,561055	2,894567	1,345345
2008M09	1025,446	2,822627	1,306703	1,317759
2008M10	1026,754	3,48053	2,17246	1,30314
2008M11	1031,027	2,109525	2,011777	1,302478
2008M12	1031,876	1,752485	2,222908	1,338369
2009M01	1032,891	2,854794	1,03622	1,328017
2009M02	1033,242	2,413409	2,744248	1,332493
2009M03	1029,904	3,334624	2,864237	1,332164
2009M04	1030,351	1,310939	2,130266	1,355932
2009M05	1030,972	4,897579	2,393302	1,406471
2009M06	1036,398	4,944843	2,844999	1,395853
2009M07	1036,514	3,792647	2,414477	1,386064
2009M08	1037,617	3,144385	1,305078	1,406948
2009M09	1038,047	2,23811	2,152577	1,420586
2009M10	1036,21	4,25518	2,21343	1,45752
2009M11	1038,996	3,738925	1,848261	1,469198
2009M12	1040,999	1,650468	2,472888	1,462012
2010M01	1043,081	4,643709	2,868734	1,473826
2010M02	1041,763	4,290149	2,851137	1,496

2010M03	1045,068	4,7992	1,901679	1,512409
2010M04	1042,765	3,902878	1,226476	1,522555
2010M05	1044,438	3,453661	2,969682	1,543888
2010M06	1049,319	2,672972	2,677796	1,567274
2010M07	1047,838	4,730914	1,249325	1,594917
2010M08	1047,206	4,464256	2,841684	1,607891
2010M09	1047,905	1,180875	2,739793	1,604549
2010M10	1047,398	1,105468	2,037676	1,607483
2010M11	1044,797	2,505853	2,182551	1,631614
2010M12	1045,434	4,242213	1,798005	1,615275
2011M01	1043,809	4,949105	1,109523	1,622648
2011M02	1045,256	1,601668	1,670394	1,614782
2011M03	1043,917	3,376523	2,605707	1,615356
2011M04	1047,517	2,523563	1,009264	1,640925
2011M05	1046,451	4,879658	1,666998	1,644747
2011M06	1046,307	4,368476	1,796337	1,645676
2011M07	1048,434	4,353315	2,074791	1,618479
2011M08	1046,472	2,874773	2,839711	1,633404
2011M09	1047,427	2,659278	1,692692	1,646314
2011M10	1050,541	2,093628	1,693906	1,689579
2011M11	1047,826	1,225502	2,475002	1,683423
2011M12	1048,695	4,45889	1,904436	1,687806
2012M01	1049,715	4,251604	1,44921	1,692794
2012M02	1051,779	4,998871	1,904879	1,724343
2012M03	1049,805	4,986547	1,281714	1,722437
2012M04	1047,664	3,221727	1,352774	1,728018
2012M05	1049,208	4,07595	1,996736	1,740176
2012M06	1050,302	4,779063	1,837851	1,743908
2012M07	1051,303	4,39859	2,829692	1,734979
2012M08	1052,496	1,989392	1,724788	1,738861
2012M09	1051,636	2,802177	2,161177	1,760334
2012M10	1052,6	1,516638	2,264529	1,739803
2012M11	1053,686	4,816204	1,026189	1,742463
2012M12	1052,758	3,424699	2,327075	1,72846
2013M01	1056,989	1,914571	1,356072	1,752361
2013M02	1058,437	3,686803	2,922141	1,721897
2013M03	1056,554	3,472513	1,297325	1,710719
2013M04	1058,367	2,432651	1,829248	1,718263
2013M05	1056,918	1,45423	1,170699	1,749574
2013M06	1058,992	3,686293	2,993749	1,748259
2013M07	1061,809	3,081231	2,00439	1,737155
2013M08	1060,668	4,089274	2,19077	1,774778
2013M09	1063,095	3,080654	1,134153	1,745818
2013M10	1064,42	4,408726	2,499921	1,701841
2013M11	1066,565	3,207627	1,419811	1,710642
2013M12	1070,858	3,243752	2,796109	1,700601

2014M01	1070,867	4,506614	1,410279	1,680176
2014M02	1069,86	2,613931	1,381375	1,694343
2014M03	1068,581	1,536061	1,073099	1,699219
2014M04	1067,449	1,115131	1,944134	1,687938
2014M05	1067,795	4,020549	2,129682	1,662331
2014M06	1068,977	3,481238	1,131417	1,679781
2014M07	1070,031	3,816319	2,551055	1,692785
2014M08	1072,185	1,851857	1,906578	1,690801
2014M09	1072,711	1,545486	2,048781	1,727734
2014M10	1076,118	1,058179	1,881525	1,706332
2014M11	1076,089	2,40235	1,801526	1,675822
2014M12	1082,029	3,359671	2,119281	1,661983
2015M01	1083,78	2,568976	1,31048	1,661072
2015M02	1082,566	2,7499	1,363856	1,665939
2015M03	1080,924	4,616635	2,723571	1,661114
2015M04	1082,389	2,393022	2,892231	1,668155
2015M05	1082,442	3,055958	1,746619	1,643124
2015M06	1084,37	4,134612	1,541489	1,671999
2015M07	1085,817	2,586171	2,287999	1,670356
2015M08	1086,171	3,488347	1,817468	1,692702
2015M09	1084,978	4,449455	1,050773	1,699557
2015M10	1082,448	4,798082	1,312305	1,708692
2015M11	1082,055	1,588294	2,431944	1,720087
2015M12	1084,268	4,706351	2,317848	1,729041
2016M01	1085,196	2,968465	1,054192	1,741896
2016M02	1083,204	2,032978	1,443944	1,768479
2016M03	1084,051	2,836543	1,46215	1,772409
2016M04	1085,321	4,92013	2,343785	1,786589
2016M05	1084,054	2,970472	1,039421	1,784795
2016M06	1084,861	2,315006	1,208217	1,813597
2016M07	1085,478	3,533603	2,599832	1,800069
2016M08	1083,692	1,960582	1,357089	1,836088
2016M09	1084,907	1,303453	2,305492	1,835285
2016M10	1086,529	1,515519	1,476366	1,806669
2016M11	1089,195	1,512183	1,198883	1,809231
2016M12	1091,802	1,607611	1,486344	1,79561
2017M01	1089,547	1,555309	2,444534	1,812423
2017M02	1088,171	3,563499	2,711393	1,799371
2017M03	1089,702	1,72752	2,66044	1,790447
2017M04	1091,229	2,382669	1,794367	1,752656
2017M05	1092,759	4,587154	2,33617	1,74361
2017M06	1100,965	2,895847	1,409969	1,695133
2017M07	1102,606	3,670231	1,586295	1,663455
2017M08	1105,378	1,689279	2,792672	1,678663
2017M09	1107,786	1,769156	1,026004	1,694379
2017M10	1109,588	1,163474	1,171017	1,702888

2017M11	1109,458	1,67574	1,415773	1,683548
2017M12	1111,476	2,114361	1,053064	1,682594
2018M01	1110,43	1,708042	1,362871	1,682522
2018M02	1110,456	1,35481	2,166083	1,659355
2018M03	1109,986	1,482543	1,842849	1,689423
2018M04	1110,65	2,843115	2,785343	1,70697
2018M05	1115,779	1,825335	2,634887	1,702551
2018M06	1112,544	2,457079	1,683635	1,703089
2018M07	1114,417	3,013669	1,518847	1,707256
2018M08	1111,691	3,761579	1,759385	1,666421
2018M09	1111,248	1,157249	2,18059	1,661478
2018M10	1113,925	4,197642	1,536127	1,647838
2018M11	1114,554	3,511602	2,248298	1,627806
2018M12	1112,898	1,327036	1,818823	1,622184
2019M01	1111,968	4,494314	2,104094	1,658138
2019M02	1113,827	4,68349	1,872253	1,670954
2019M03	1112,866	1,244312	1,588932	1,659531
2019M04	1113,799	2,107511	2,896907	1,670983
2019M05	1114,39	4,224805	2,527212	1,69897
2019M06	1113,587	3,993039	1,280226	1,717462
2019M07	1118,375	1,738084	2,736936	1,718655
2019M08	1120,143	1,837397	1,974862	1,705716
2019M09	1116,593	2,481888	2,789104	1,719681
2019M10	1117,466	2,938092	2,599711	1,72755
2019M11	1116,642	3,473019	1,850427	1,745454
2019M12	1118,847	2,475655	1,044939	1,758158
2020M01	1117,762	2,850139	1,537355	1,779149
2020M02	1118,032	3,989884	2,083268	1,768444
2020M03	1119,542	1,146733	2,266956	1,794792
2020M04	1121,774	2,009748	1,515775	1,798744
2020M05	1119,873	3,853398	1,278712	1,840249
2020M06	1119,704	4,580827	2,66986	1,826465
2020M07	1119,254	3,04671	2,968804	1,861185
2020M08	1118,448	3,128454	2,05138	1,865143
2020M09	1122,479	1,428688	1,343359	1,852114
2020M10	1123,789	2,789649	1,544615	1,842437
2020M11	1121,767	3,130469	1,036781	1,83603
2020M12	1124,103	1,969882	2,828598	1,844513
2021M01	1128,847	2,076973	1,235502	1,85497
2021M02	1131,412	2,509137	2,153033	1,843496
2021M03	1128,873	1,080285	1,54811	1,843009
2021M04	1128,405	2,288317	2,108356	1,885854
2021M05	1131,438	1,845792	2,302841	1,920405
2021M06	1130,523	2,309989	2,659484	1,929131
2021M07	1131,911	1,479049	1,412843	1,929892
2021M08	1133,96	4,562109	1,021992	1,932292

2021M09	1132,606	3,37437	1,273771	1,944563
2021M10	1132,987	3,716409	2,800037	1,924107
2021M11	1127,004	4,156685	2,74778	1,918959
2021M12	1125,456	2,993769	2,194826	1,885587
2022M01	1125,451	1,347681	2,201034	1,893572
2022M02	1123,455	3,148426	2,330073	1,906516
2022M03	1127,22	3,347364	1,350743	1,896852
2022M04	1124,86	3,981758	2,828824	1,928332
2022M05	1124,479	2,726638	1,837541	1,903817
2022M06	1125,241	1,510321	1,766277	1,874529
2022M07	1128,623	2,135104	2,037835	1,879018
2022M08	1126,252	2,452329	1,093932	1,89996
2022M09	1129,078	3,583669	1,332567	1,933639
2022M10	1129,599	3,283113	2,476067	1,924461
2022M11	1128,136	2,424387	1,165597	1,946035
2022M12	1129,56	4,946061	2,206304	1,945264
2023M01	1130,458	3,423099	1,490698	1,941812
2023M02	1129,757	1,948907	1,778591	1,959485
2023M03	1130,397	1,40713	1,577387	1,972531
2023M04	1130,126	1,611437	1,711345	1,941004
2023M05	1130,853	1,983831	2,438092	1,970534
2023M06	1132,678	1,642725	1,594243	1,998136
2023M07	1136,35	1,746268	2,132809	1,985625
2023M08	1134,374	2,140381	1,952101	1,993541
2023M09	1139,14	1,693494	2,327342	2,003422
2023M10	1135,736	4,587062	2,873659	2,008635
2023M11	1135,932	1,320935	2,465144	1,997629
2023M12	1137,609	3,098046	1,429881	1,984197
2024M01	1138,671	2,641587	1,062366	1,983686
2024M02	1137,926	4,929514	1,524528	2,00714
2024M03	1138,009	1,448156	2,190156	2,018012
2024M04	1137,523	2,591422	1,102852	2,0106
2024M05	1136,845	4,877882	1,992732	2,026034
2024M06	1139,044	4,462029	2,193686	1,969063
2024M07	1140,258	4,268288	1,668488	1,992038
2024M08	1139,372	2,031611	2,541824	1,957244
2024M09	1141,671	1,68355	1,213197	1,949995
2024M10	1142,786	3,674573	1,150276	1,927602
2024M11	1144,912	4,717504	2,456378	1,901708
2024M12	1146,671	3,227052	1,990983	1,924925

Source : Données fictives.

Estimation du modèle de base : Avant d'effectuer les tests de stabilité, nous devons estimer un modèle économétrique, supposons que nous voulons étudier la relation entre l'indice boursier et les autres variables financières :

1. Aller dans **Quick > Estimate Equation**,

- Entrer le modèle suivant : **Indice_Boursier C Taux_Interet Inflation Taux_Change**
- Cliquer sur OK pour estimer le modèle.

Tableau 32 : Estimation de l'indice boursier en fonction des autres variables.

Dependent Variable: INDICE_BOURSIER
 Method: Least Squares
 Date: 02/20/25 Time: 19:55
 Sample: 2000M01 2024M12
 Included observations: 300

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	834,8690	7,488514	111,4866	0,0000
TAUX_INTERET	-2,005938	0,856208	-2,342816	0,0198
INFLATION	-0,324121	1,814878	-0,178591	0,8584
TAUX_CHANGE	151,2767	3,595799	42,07041	0,0000
R-squared	0,860479	Mean dependent var		1063,302
Adjusted R-squared	0,859065	S,D, dependent var		46,96016
S,E, of regression	17,62948	Akaike info criterion		8,590266
Sum squared resid	91996,37	Schwarz criterion		8,639649
Log likelihood	-1284,540	Hannan-Quinn criter,		8,610029
F-statistic	608,5135	Durbin-Watson stat		0,076586
Prob(F-statistic)	0,000000			

Source : réalisé par l'auteur en utilisant Eviews 10.

A/ Test de Chow (Rupture structurelle à une date précise) : Le test de Chow permet de vérifier si les coefficients du modèle changent après une certaine période.

- Sélectionner la **régression estimée**.
- Aller dans **View** → **Stability Tests** → **Chow Breakpoint Test**.
- Spécifier une date de rupture (exemple : 2010M01 si on suspecte une rupture en janvier 2010).
- Cliquer sur **OK** et analyser le résultat :

Tableau 33 : Le test de Chow.

Chow Breakpoint Test: 2010M01
 Null Hypothesis: No breaks at specified breakpoints
 Varying regressors: All equation variables
 Equation Sample: 2000M01 2024M12

F-statistic	7,652625	Prob, F(4,292)	0,0000
Log likelihood ratio	29,90757	Prob, Chi-Square(4)	0,0000
Wald Statistic	30,61050	Prob, Chi-Square(4)	0,0000

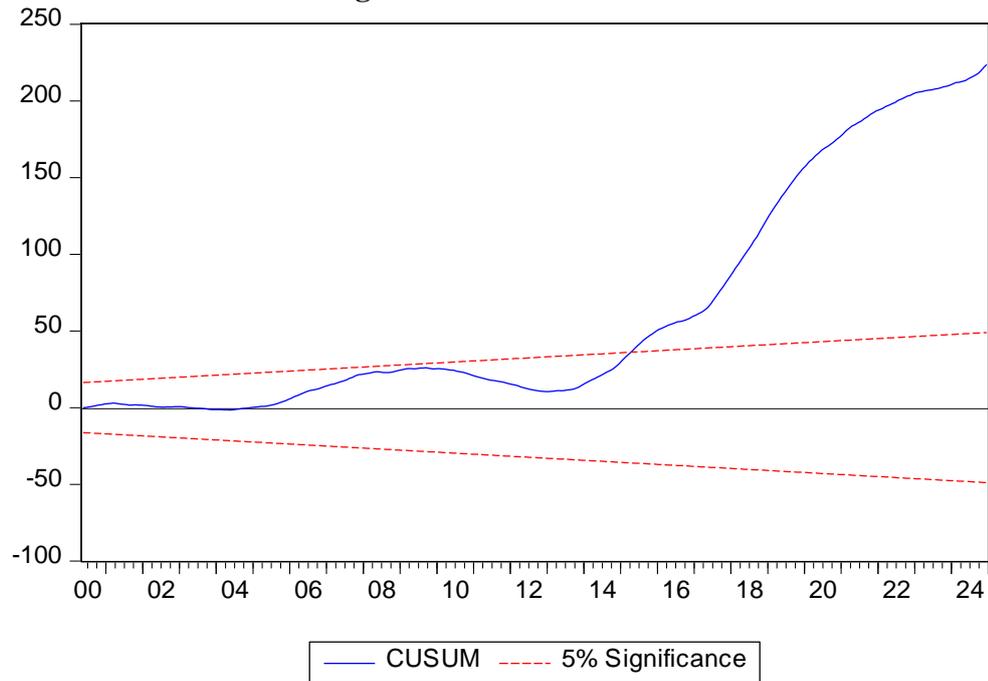
Source : réalisé par l'auteur en utilisant Eviews 10.

La statistique F est significative (p-value < 5%), il y a donc une rupture structurelle.

B/ Tests CUSUM et CUSUMSQ (Stabilité des résidus) : Ces tests permettent de vérifier la stabilité des coefficients sur toute la période.

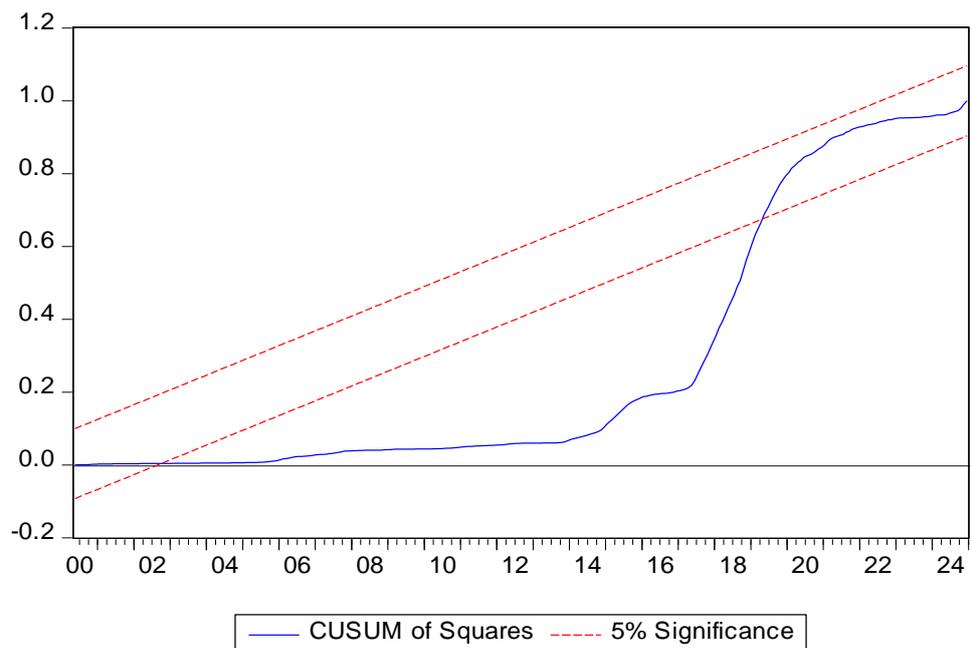
1. Sélectionner la **régression estimée**.
2. Aller dans **View** → **Stability Tests** → **CUSUM Test** ou **CUSUM of Squares Test**.
3. Examiner les graphiques :

Figure 13 : Test CUSUM.



Source : réalisée par l'auteur en utilisant Eviews 10.

Figure 14 : Test CUSUMSQ.



Source : réalisée par l'auteur en utilisant Eviews 10.

La courbe, dans les deux tests CUSUM et CUSUM of Squares, sort des bandes, il y a instabilité structurelle.

C/ Test de Bai-Perron (Multiples points de rupture) : Si l'on suspecte plusieurs ruptures, le test de Bai-Perron est plus adapté.

1. Sélectionner la **régression estimée**.
2. Aller dans **View** → **Stability Tests** → **Multiple Breakpoint Test**.
3. Définir le nombre maximum de points de rupture.
4. Cliquer sur **OK** et analyser les dates de rupture détectées.

Tableau 34 : Le test de C Test de Bai-Perron.

Multiple breakpoint tests
 Bai-Perron tests of L+1 vs, L sequentially determined breaks
 Date: 02/20/25 Time: 20:10
 Sample: 2000M01 2024M12
 Included observations: 300
 Breaking variables: C TAUX_INTERET INFLATION
 TAUX_CHANGE
 Break test options: Trimming 0,15, Max, breaks 5, Sig, level 0,05

Sequential F-statistic determined breaks: 3

Break Test	F-statistic	Scaled F-statistic	Critical Value**
0 vs, 1 *	233,3588	933,4351	16,19
1 vs, 2 *	92,53989	370,1596	18,11
2 vs, 3 *	47,44191	189,7676	18,93
3 vs, 4	1,609641	6,438566	19,64

* Significant at the 0,05 level,
 ** Bai-Perron (Econometric Journal, 2003) critical values,

Break dates:

	Sequential	Repartition
1	2017M06	2005M08
2	2013M09	2013M09
3	2005M08	2017M06

Source : réalisé par l'auteur en utilisant Eviews 10.

Le test de Bai-Perron est utilisé pour identifier les ruptures structurelles dans un modèle. Il compare les coefficients du modèle sur différentes périodes pour voir s'ils changent significativement.

Nombre de ruptures détectées : "Sequential F-statistic determined breaks: 3" signifie que trois ruptures structurelles ont été détectées dans la relation entre l'indice boursier et les variables explicatives (taux d'intérêt, inflation, taux de change), Cela suggère que les coefficients du modèle changent à trois moments différents dans la période 2000-2024.

Significativité des ruptures : Le test compare différents modèles avec un nombre croissant de ruptures :

- **0 vs, 1 rupture** → F-statistic = **233,36** ($p < 0,05$, donc significatif)
- **1 vs, 2 ruptures** → F-statistic = **92,54** ($p < 0,05$, donc significatif)
- **2 vs, 3 ruptures** → F-statistic = **47,44** ($p < 0,05$, donc significatif)
- **3 vs, 4 ruptures** → F-statistic = **1,61** ($p > 0,05$, donc non significatif)

En somme, les trois (3) premières ruptures sont statistiquement significatives, mais la quatrième rupture ne l'est pas.

Dates des ruptures détectées : Le tableau donne deux ensembles de dates séquentielle (méthode séquentielle) et répartition (méthode alternative). Les dates des ruptures :

- **1ère rupture** : Juin 2017 (Méthode séquentielle) ou Août 2005 (Méthode de répartition)
- **2ème rupture** : Septembre 2013 (confirmée par les deux méthodes)
- **3ème rupture** : Août 2005 (Méthode séquentielle) ou Juin 2017 (Méthode de répartition)

Interprétation possible :

- **2005-2006** : Crise financière mondiale naissante (effondrement des marchés émergents).
- **2013** : Crise de la dette européenne et changement de politique monétaire aux États-Unis.
- **2017** : Montée des taux d'intérêt et changements dans les politiques économiques globales.

Conclusion

Les tests de stabilité sont essentiels pour vérifier la robustesse d'un modèle économétrique. Selon le type d'instabilité suspectée (brusque ou graduelle), le choix du test approprié est décisif. L'utilisation combinée du test de Chow, des tests CUSUM et CUSUMSQ ainsi que du test de Bai-Perron permet d'assurer une détection complète des ruptures structurelles. Les applications sous Eviews facilitent leur mise en œuvre et l'interprétation des résultats.

Chapitre 10 : La méthodologie de Box-Jenkins

Introduction

La méthodologie de Box-Jenkins²⁶ est une approche statistique pour la modélisation des séries temporelles, basée sur les modèles ARIMA (Auto Regressive Integrated Moving Average). Elle permet de capturer les dynamiques temporelles et d'effectuer des prévisions fiables.

1) Les étapes de la méthodologie Box-Jenkins

Un processus autorégressif, noté AR, est caractérisé par une combinaison linéaire finie de ses propres valeurs passées, De son côté, un processus à moyenne mobile, noté MA, est défini par une combinaison linéaire des valeurs passées d'un bruit blanc. Selon Wold (1954), les modèles ARMA offrent une représentation efficace de la plupart des processus stationnaires. La méthodologie proposée par Box et Jenkins (1976) vise à analyser systématiquement les séries chronologiques en fonction de leurs propriétés afin d'identifier, parmi les modèles ARIMA, celui qui décrit le mieux le phénomène étudié. Cette approche repose sur plusieurs étapes :

a) Identification du modèle :

L'objectif est de déterminer la structure ARIMA la plus adaptée à la série temporelle. Le modèle ARIMA, basé sur la méthodologie Box-Jenkins, requiert que la série soit stationnaire. Par ailleurs, une série avec une saisonnalité n'est pas stationnaire, car ses moyennes et variances changent selon les saisons. Dans le cas d'une série non désaisonnalisée, les coefficients du modèle peuvent capter artificiellement des effets saisonniers au lieu des vraies dynamiques de la série pour rendre le modèle moins fiable et son interprétation aussi.

- **Analyse de la stationnarité** : Vérifier si la série est stationnaire à l'aide du test de Dickey-Fuller Augmenté (ADF).
- **Correlogramme** : Examiner les fonctions d'autocorrélation (ACF) et d'autocorrélation partielle (PACF) pour identifier les ordres AR et MA.

b) Estimation du modèle

Une fois l'ordre du modèle ARIMA(p,d,q) déterminé, l'étape suivante consiste à estimer ses paramètres. Cette estimation est généralement effectuée à l'aide de la méthode du maximum de vraisemblance, qui vise à trouver les valeurs des paramètres maximisant la

²⁶ Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time Series Analysis: Forecasting and Control (5th ed.). Wiley.

probabilité d'observer les données disponibles. Concrètement, cette approche repose sur la définition d'une fonction de vraisemblance, construite à partir de la distribution des erreurs du modèle, puis sur son optimisation pour identifier les paramètres les plus plausibles. Des algorithmes numériques, tels que l'algorithme de Newton-Raphson ou l'algorithme d'optimisation de Broyden-Fletcher-Goldfarb-Shanno (BFGS), sont souvent utilisés pour résoudre ce problème d'optimisation. Une fois les paramètres estimés, leur significativité statistique peut être évaluée à l'aide d'outils tels que les écarts-types et les tests de Student.

c) Diagnostic du modèle

Une fois le modèle ARIMA estimé, il est essentiel d'évaluer sa pertinence et la qualité de son ajustement aux données. Cette étape repose sur plusieurs tests et vérifications :

- ✓ Absence d'autocorrélation des résidus : Les résidus du modèle doivent se comporter comme un bruit blanc, c'est-à-dire être indépendants et non autocorrélés. Cette propriété est vérifiée à l'aide du test de Ljung-Box, qui examine si les autocorrélations des résidus sont statistiquement significatives. Une autocorrélation résiduelle indiquerait que le modèle ne capte pas pleinement la structure temporelle des données et nécessiterait une révision.
- ✓ Tests de normalité des résidus : La normalité des résidus est une hypothèse importante pour la validité de certains intervalles de confiance et tests statistiques. Elle peut être évaluée à l'aide du test de Jarque-Bera, du test de Shapiro-Wilk ou par l'observation d'un histogramme et d'un graphique quantile-quantile (Q-Q plot).
- ✓ Vérification de l'homoscédasticité : L'hypothèse d'homoscédasticité suppose que la variance des résidus reste constante dans le temps. Cette propriété peut être testée par l'observation graphique (graphe des résidus) ou via des tests statistiques comme le test de Breusch-Pagan ou le test d'Engle (ARCH) en cas de suspicion d'hétéroscédasticité conditionnelle.

Dans le cas où ces vérifications révèlent des anomalies, des ajustements du modèle peuvent être nécessaires, notamment par une modification de l'ordre ARIMA, l'inclusion de termes supplémentaires ou l'application de transformations sur la série étudiée.

d) Validation et prévision

Après avoir validé le modèle en s'assurant qu'il respecte les hypothèses statistiques (absence d'autocorrélation des résidus, normalité et homoscedasticité), il peut être utilisé pour réaliser des prévisions sur la série temporelle étudiée. La qualité des prévisions dépend directement de la précision du modèle estimé. Pour évaluer cette précision, différentes mesures d'erreur peuvent être utilisées, telles que l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne (MAE) ou encore l'erreur de prévision moyenne en pourcentage (MAPE).

Les prévisions peuvent être ponctuelles, fournissant une valeur estimée pour chaque période future, ou accompagnées d'intervalles de confiance, permettant d'évaluer l'incertitude associée aux prévisions. Ces intervalles sont généralement construits sous l'hypothèse d'une distribution normale des erreurs du modèle.

Enfin, dans le cadre d'une application pratique, il est recommandé de comparer les performances du modèle ARIMA avec d'autres méthodes de prévision (exemple : modèles exponentiels, réseaux neuronaux,...etc.) afin de s'assurer de la robustesse des résultats obtenus²⁷.

2) Application sous Eviews

a) Importation des données

1. Ouvrir Eviews et importer la base de données (ex, fichier Excel contenant une série temporelle).
2. Transformer la variable en série temporelle si nécessaire.

b) Test de stationnarité

- Aller dans **View** → **Unit Root Test** et choisir **ADF**.
- Si la série n'est pas stationnaire, appliquer une différenciation.

c) Identification du modèle

- Afficher le correlogramme (**View** → **Correlogram**) pour identifier les ordres AR et MA.

d) Estimation du modèle ARIMA

1. Aller dans **Quick** → **Estimate Equation**.
2. Entrer l'équation sous la forme :
series c ar(1) ma(1)
3. Analyser les coefficients et la significativité des paramètres.

²⁷ Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice (2nd ed.). OTexts.

NB/ Quand on estime un modèle ARIMA, l'hypothèse principale est que les résidus (les erreurs entre les valeurs prédites et les valeurs réelles) ne doivent pas être autocorrélés. Si les résidus montrent des corrélations significatives, cela signifie que le modèle n'a pas capté toute la structure de la série temporelle, et il faut alors ajuster le modèle.

e) Diagnostic du modèle

- Vérifier l'autocorrélation des résidus (**View** → **Residual Tests** → **Correlogram-Q-Statistic**).
- Tester la normalité (**View** → **Residual Tests** → **Histogram-Normality Test**).

f) Prévion avec le modèle ARIMA

1. Aller dans **Forecast** et choisir une période de prévision.
2. Examiner les intervalles de confiance et comparer avec les valeurs réelles.

Exemple d'application : Nous allons utiliser les données de l'exemple d'application du septième chapitre pour mettre en pratique la méthodologie de Box-Jenkins. Plus précisément, nous allons nous intéresser à la série désaisonnalisée.

Application de la méthodologie Box-Jenkins (modèle ARIMA) sur la variable "serie_desaison".

a) Identification du modèle :

Test de Stationnarité : Selon le test ADF (voir le chapitre 8) la série désaisonnalisée n'est pas stationnaire en niveau, mais elle est stationnaire en premières différences.

Correlogramme : l'examen des fonctions d'autocorrélation (ACF) et d'autocorrélation partielle (PACF) pour identifier les ordres AR (AutoRegressif) et MA (Moyenne Mobile) effectué dans le huitième chapitre suggère un AR(4) et MA(0) de la série différenciée.

Donc, selon la méthodologie de de Box-Jenkins le modèle ARIMA (4,1,0) semble bien approprié.

b) Estimation du modèle ARIMA (4,1,0) :

Tableau 35 : Estimation du modèle ARIMA (4,1,0).

Dependent Variable: D(SERIE_DESAISON)
 Method: ARMA Maximum Likelihood (OPG - BHHH)
 Date: 02/25/25 Time: 21:01
 Sample: 2001Q1 2015Q4
 Included observations: 60
 Convergence achieved after 9 iterations
 Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std, Error	t-Statistic	Prob,
----------	-------------	------------	-------------	-------

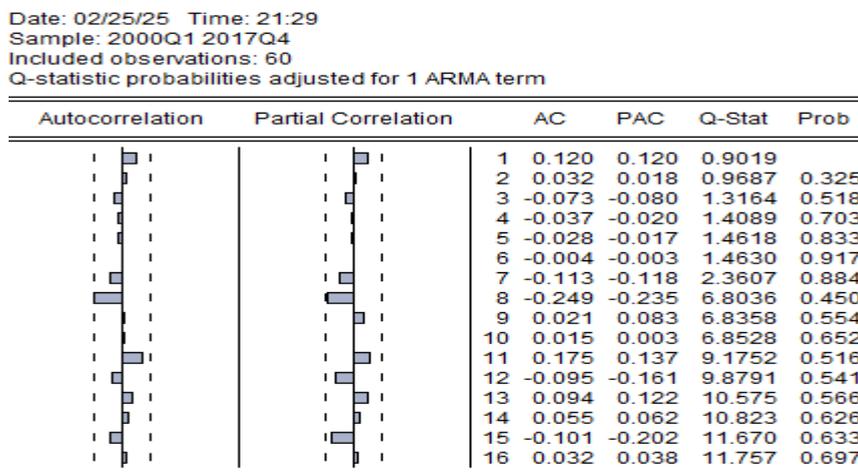
C	7,150128	0,589728	12,12444	0,0000
AR(4)	-0,444506	0,135908	-3,270652	0,0018
SIGMASQ	39,13954	9,566076	4,091494	0,0001
R-squared	0,181614	Mean dependent var		7,343933
Adjusted R-squared	0,152899	S,D, dependent var		6,973943
S,E, of regression	6,418685	Akaike info criterion		6,619686
Sum squared resid	2348,373	Schwarz criterion		6,724403
Log likelihood	-195,5906	Hannan-Quinn criter,		6,660646
F-statistic	6,324645	Durbin-Watson stat		1,725043
Prob(F-statistic)	0,003306			
Inverted AR Roots	,58+,58i	,58+,58i	-,58-,58i	-,58-,58i

Source : réalisé par l'auteur en utilisant Eviews 10.

Interprétation :

- **C (constante) = 7,15** : La moyenne du changement de la série (tendance) est significative.
 - **AR(4) = -0,44 (p = 0,0018)** : Le coefficient est significatif, ce qui signifie que la valeur de la série dépend de sa valeur quatre périodes auparavant.
 - **SIGMASQ = 39,14** : Variance de l'erreur, importante pour l'ajustement du modèle.
 - **R² = 0,18 et R² ajusté = 0,15** : Très faible, ce qui signifie que ce modèle explique peu la variabilité de la série.
 - **S,E, de la régression = 6,42** : Indique l'écart moyen entre les valeurs ajustées et réelles.
 - **Durbin-Watson = 1,73** : Correct, pas de forte autocorrélation des résidus mais mérite d'être confirmé avec des tests spécifiques.
- c) **Diagnostic du modèle** : Test d'autocorrélation et test de normalité des résidus
- Vérification de l'autocorrélation des résidus

Figure 15 : Test d'autocorrélation.

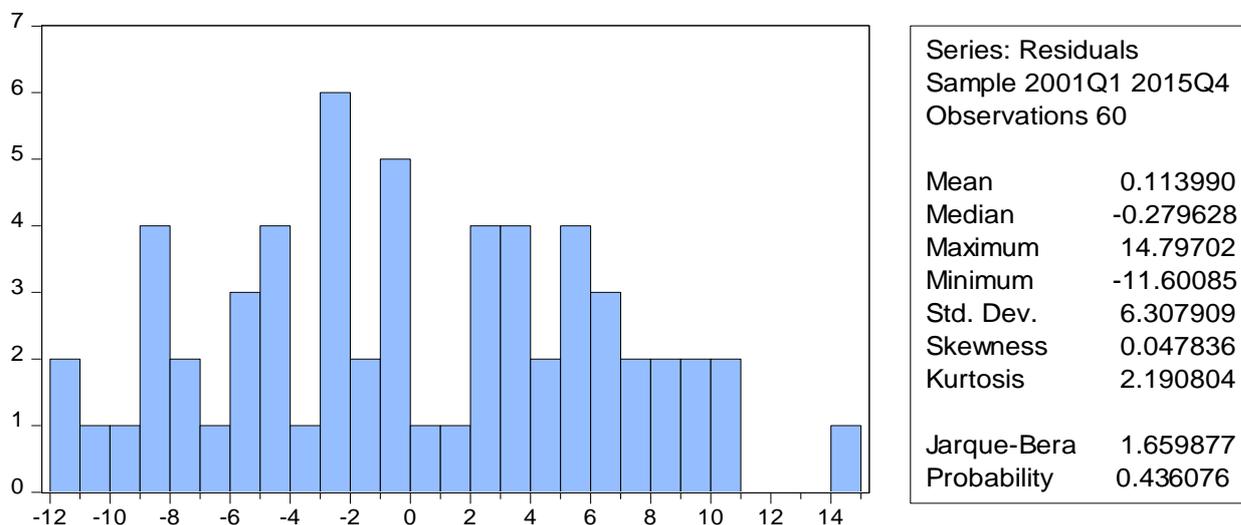


Source : réalisée par l'auteur en utilisant Eviews 10.

Le modèle ARIMA(4,1,0) est bien spécifié car les erreurs ne présentent pas d'autocorrélation significative. Cela valide l'hypothèse que les résidus suivent un bruit blanc, ce qui est une condition essentielle pour un bon modèle de prévision.

- Test de la normalité

Figure 16 : Test de normalité des erreurs.



Source : réalisée par l'auteur en utilisant Eviews 10.

Les résidus suivent une distribution normale, ce qui permet de valider le modèle.

Tableau 36 : Test d'hétéroscédasticité.

Heteroskedasticity Test: ARCH

F-statistic	0,618962	Prob, F(1,57)	0,4347
Obs*R-squared	0,633798	Prob, Chi-Square(1)	0,4260

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 02/26/25 Time: 21:23

Sample (adjusted): 2001Q2 2015Q4

Included observations: 59 after adjustments

Variable	Coefficient	Std, Error	t-Statistic	Prob,
C	34,26524	7,684629	4,458932	0,0000
RESID^2(-1)	0,103393	0,131419	0,786741	0,4347

R-squared	0,010742	Mean dependent var	38,37966
Adjusted R-squared	-0,006613	S,D, dependent var	43,10722
S,E, of regression	43,24952	Akaike info criterion	10,40516
Sum squared resid	106619,7	Schwarz criterion	10,47558
Log likelihood	-304,9522	Hannan-Quinn criter,	10,43265
F-statistic	0,618962	Durbin-Watson stat	1,856277
Prob(F-statistic)	0,434692		

Source : réalisé par l'auteur en utilisant Eviews 10.

d) **Prévisions futures** (Prévision (Forecast) sur une période de 8 trimestres, examen des intervalles de confiance et comparaison avec les valeurs réelles). Le modèle ARIMA (4,1,0) est estimé sur la série désaisonnalisée ("serie_desaison"). Il produit donc des prévisions qui n'intègrent pas encore la saisonnalité. Donc, il faut utiliser le modèle ARIMA (4,1,0) pour générer des prévisions sur les périodes futures. Les prévisions sont obtenues sur la série désaisonnalisée, il faut réintroduire la saisonnalité pour revenir à la série originale. Le modèle initial est additif, donc :

Prévisions corrigées = prévisions désaisonnalisées + Composante saisonnière

Cependant, dans le cas d'un modèle multiplicatif la méthode de désaisonnalisation utilisée est différente. D'où :

Prévisions corrigées = prévisions désaisonnalisées × Composante saisonnière

Dans Eviews, la procédure pour tenir compte de la saisonnalité après l'estimation d'un modèle ARIMA sur une série désaisonnalisée est la suivante :

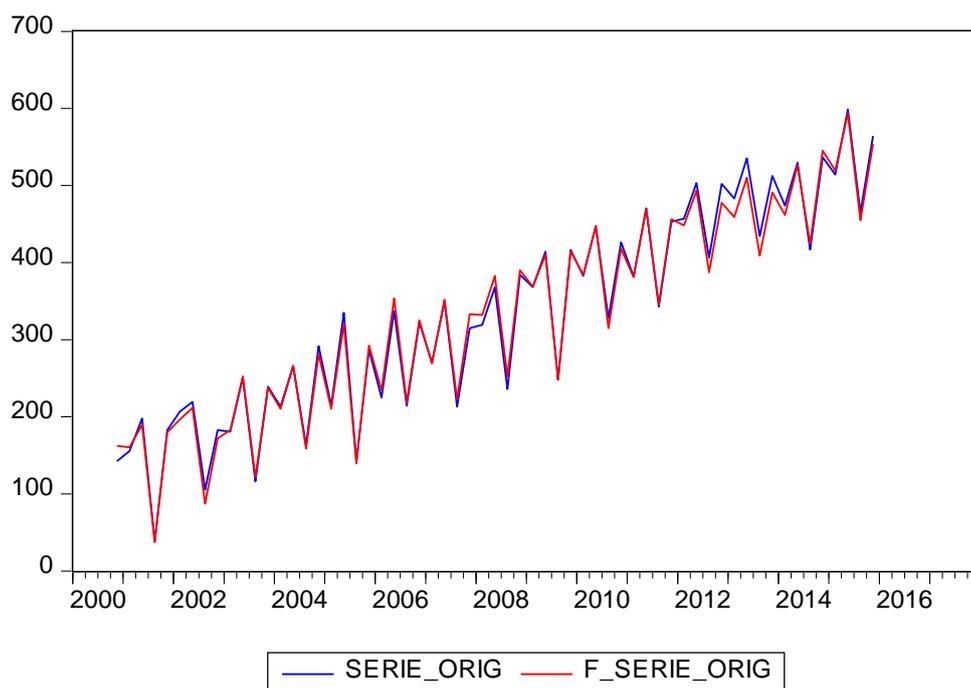
- 1) Aller dans le menu **Proc** de la fenêtre du modèle estimé.
- 2) Sélectionner **Forecast** (Prévisions).
- 3) Dans **Forecast name**, donner un nom à la série de prévisions, par exemple **F_SERIE_DESAISON**.
- 4) Cliquer sur **OK** → Eviews génère une série de prévisions **sans saisonnalité**.

Pour réintroduire la saisonnalité, on ajoute la composante saisonnière. D'où, la série générée des prévisions corrigées :

SERIE_ORIG = SERIE_DESAISON + COMPOSANTE_SAISONNIERE

D'où, **F_SERIE_ORIG = F_SERIE_DESAISON + COMPOSANTE_SAISONNIERE**

Figure 17 : Séries de prévisions avec saisonnalité.



Source : réalisée par l'auteur en utilisant Eviews 10.

Les prévisions semblent suivre étroitement la série originale, ce qui indique que le modèle capte bien la tendance et la saisonnalité :

- Il n'y a pas d'écart systématique visible (pas de sous-estimation ou surestimation constante), donc le modèle n'est pas biaisé.
- Les pics et creux saisonniers sont bien reproduits, ce qui confirme que la saisonnalité a été correctement réintroduite après la désaisonnalisation.

Conclusion

La méthodologie de Box-Jenkins est un outil puissant pour l'analyse et la prévision des séries temporelles. Son application sous Eviews permet d'obtenir des modèles prédictifs fiables et adaptés aux données économiques et financières.

Conclusion Générale

L'étude des techniques d'analyse économétrique et de modélisation statistique à travers le logiciel Eviews constitue une opportunité pour les étudiants et leur offre un guide à la fois théorique et pratique permettant d'analyser des données économiques et financières de manière rigoureuse et pertinente. Ce cours a permis d'explorer plusieurs thématiques fondamentales allant de la régression linéaire simple et multiple à des concepts avancés tels que la détection et la correction des problèmes économétriques (corrélation des erreurs, multicollinéarité, hétéroscédasticité), en passant par les tests aléatoires et les méthodes de désaisonnalisation.

L'application de la méthodologie de Box-Jenkins et l'utilisation des fonctions d'autocorrélation (ACF) et d'autocorrélation partielle (PACF) ont permis de mieux appréhender la modélisation des séries temporelles et l'établissement des prévisions d'une série de données à partir de ses observations passées. Ces outils permettent non seulement de mieux comprendre le comportement des variables étudiées, mais aussi d'aider à la prise de décision stratégique basée sur une modélisation empirique solide à partir des données réelles.

Ainsi, ce cours constitue une base de connaissances nécessaires pour les étudiants de première année Master Comptabilité et Audit et des autres niveau ou spécialités souhaitant développer des compétences excellentes dans l'analyse de données économétriques et statistiques. La maîtrise des techniques détaillées dans ce cours ouvre la voie à une application pratique dans divers domaines tels que la finance, l'économie, le marketing, et la gestion des risques. En somme, l'utilisation judicieuse des outils économétriques permet de mieux comprendre les mécanismes économiques et d'anticiper les évolutions futures avec plus de précision et de fiabilité.

Références bibliographique

Abderrahmani, F. (2018). *Guide pratique des séries temporelles macro-économiques et financières avec Eviews 9.5*. Polycopié de cours à caractère pédagogique, Université de Bejaia.

Asteriou, D. & Hall, S. G. (2021). *Applied econometrics* (4th ed.). Palgrave Macmillan.

Bai, J. & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*. 18(1). 1-22. <https://doi.org/10.1002/jae.659>

Box, G. E. P., Jenkins, G. M., Reinsel, G. C. & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.

Box, G. E. P. & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*. 65(332), 1509-1526.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*. 17(31), 334–355.

Brockwell, P. J. & Davis, R. A. (2016). *Introduction to time series and forecasting* (3rd ed.), Springer.

Brooks, C. (2019). *Introductory econometrics for finance* (4th ed.). Cambridge University Press.

Buys-Ballot, C. H. D. (1847). *Les changements périodiques de température, observés d'après les moyennes de plusieurs années* (in Dutch). Utrecht.

Chatfield, C. (2003). *The analysis of time series: An introduction* (6th ed.). Chapman & Hall/CRC.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3), 591-605, <https://doi.org/10.2307/1910133>

Durbin, J. & Watson, G. S. (1950), Testing for serial correlation in least squares regression I. *Biometrika*, 37(3-4), 409–428.

Durbin, J. & Watson, G. S. (1951). Testing for serial correlation in least squares regression II. *Biometrika*, 38(1-2), 159–179.

Dufrénot, G. & Mignon, V. (2012). *Économétrie : Méthodes et applications*. Éditions Economica.

Enders, W. (2014). *Applied econometric time series* (4th ed.). Wiley.

Eviews (2023). *User's guide*, IHS Global Inc, <https://www.eviews.com>

- Farrar, D. E. & Glauber, R. R. (1967). Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics*, 49(1), 92–107.
- Greene, W. H. (2018). *Econometric analysis* (8th ed.). Pearson.
- Gujarati, D. N. & Porter, D. C. (2009). *Basic econometrics* (5th ed.). McGraw-Hill Education.
- Gujarati, D. N. & Porter, D. C. (2021). *Basic econometrics* (6th ed.). McGraw-Hill Education.
- Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.
- Hansen, B. E. (2001). The new econometrics of structural change: Dating breaks in U.S. labor productivity. *Journal of Economic Perspectives*, 15(4), 117-128, <https://doi.org/10.1257/jep.15.4.117>
- Hill, R. C., Griffiths, W. E. & Lim, G. C. (2018). *Principles of econometrics* (5th ed.). Wiley.
- Hoerl, A. E. & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67.
- Hyndman, R. J. & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.
- IHS Global Inc, (2023). *Eviews User's Guide*, IHS Markit.
- Kennedy, P. (2008). *A guide to econometrics* (6th ed.). Wiley-Blackwell.
- Kufel, T. (2011). *Econometric analysis with Eviews*. Springer.
- Lütkepohl, H. & Krätzig, M. (2004). *Applied time series econometrics*. Cambridge University Press.
- Quandt, R. E. (1960), Tests of the hypothesis that a linear regression system obeys two separate regimes, *Journal of the American Statistical Association*, 55(290), 324-330, <https://doi.org/10.2307/2281745>
- Stock, J. H. & Watson, M. W. (2019). *Introduction to econometrics* (4th ed.). Pearson.
- Stock, J. H. & Watson, M. W. (2020). *Introduction to econometrics* (4th ed.). Pearson.
- Studenmund, A. H. (2017). *Using econometrics: A practical guide* (7th ed.). Pearson.
- Tsay, R. S. (2010). *Analysis of financial time series* (3rd ed.). Wiley.
- Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (7th ed.). Cengage Learning.

Tables des matières

Introduction générale	1
Chapitre 1 : Introduction à Eviews et régression Linéaire simple	2
Introduction	2
1) Présentation du logiciel Eviews	2
1-1) Caractéristiques principales du logiciel Eviews	2
1-2) Utilisation du logiciel Eviews en analyse économique	3
2) Régression Linéaire simple dans Eviews	3
2-1) Définition et principe	3
2-2) Estimation d'un modèle de régression linéaire simple dans Eviews	3
2-2-1) Importer ou saisir les données	3
2-2-2) Définir les variables	5
2-2-3) Estimer le modèle	5
2-3) Interprétation des résultats	5
Exemple d'Application	5
Conclusion	9
Chapitre 2 : La régression multiple dans Eviews	10
Introduction	10
1) Présentation de la régression multiple	10
2) Estimation d'une régression multiple dans Eviews	10
2-1) Importation des données	10
2-2) Lancement de la régression multiple	11
3) Interprétation des résultats	11
3-1) Lecture des coefficients	11
3-2) Significativité statistique	11
3-3) Qualité globale du modèle	11
4) Test de validité du modèle	11
4-1) Test de multicollinéarité	11
4-2) Test d'hétéroscédasticité	12
4,3) Test d'autocorrélation	12
Exemple d'application d'une régression multiple dans Eviews	12
Conclusion	16
Chapitre 3 : Corrélacion des erreurs	17
Introduction	17
1) Tests de détection de la corrélation des erreurs	17
1-1) Test de Durbin-Watson	17
1-2) Test de Breusch-Godfrey	17
1-3) Test de Ljung-Box	17
2) Résolution des problèmes de corrélation des erreurs	18
2-1) Transformation de Cochrane-Orcutt	18
2-2) Modélisation avec les erreurs ARMA	18
2-3) Utilisation des moindres carrés généralisés (GLS)	18

Exemple d'application (Détection et correction de l'autocorrélation dans une régression financière) :	18
Conclusion	24
Chapitre 4 : Multiplicité linéaire (Multicolinéarité)	25
Introduction	25
1) Détection de la multiplicité linéaire	25
1-1) Analyse des corrélations croisées	25
1-1-1) Facteur d'inflation de la variance (Variance Inflation Factor – VIF)	25
1-1-2) Test de Farrar-Glauber	26
1-2) Analyse des coefficients instables	27
2) Résolution des problèmes de multiplicité linéaire	27
2-1) Élimination de variables explicatives	27
2-2) Transformation des variables	27
2-3) Combinaison linéaire des variables	27
2-4) Régression Ridge (Ridge Regression)	27
Exemple d'application	28
Conclusion	30
Chapitre 5 : Variance de différence (Hétéroscédasticité)	31
Introduction	31
1) Détection de l'hétéroscédasticité	31
1-1) Analyse graphique	31
1-2) Test de Breusch-Pagan	31
1-3) Test de White	31
1-4) Test de Goldfeld-Quandt	31
2) Résolution des problèmes de variance de différence	32
2-1) Transformation des données	32
2-2) Modèle des moindres carrés pondérés (WLS)	32
2-3) Utilisation des erreurs standards robustes (White)	32
Exemple d'application sous Eviews	32
Conclusion	37
Chapitre 6 : Tests aléatoires	38
Introduction	38
1) Test de normalité des erreurs (Jarque-Bera)	38
1-1) Test et calcul de la statistique de Jarque-Bera	38
1-2) Correction en cas de non-normalité	39
2) Test d'autocorrélation (Durbin-Watson)	39
2-1) Test et interprétation de la statistique de Durbin-Watson	39
2-2) Correction de l'autocorrélation	40
3) Test d'hétéroscédasticité (Breusch-Pagan, White)	40
3-1) Test et interprétation des statistiques de Breusch-Pagan et White	40
3-2) Correction de l'hétéroscédasticité	41
4) Test de multicolinéarité (Variance Inflation Factor - VIF)	41
4-1) Test et statistique VIF	41
4-2) Correction de la multicolinéarité	41

Exemple d'application sous Eviews	42
Conclusion	44
Chapitre 7 : Détection de la Saisonnalité et méthodes de désaisonnalisation dans Eviews	45
Introduction	45
1) Détection de la saisonnalité	45
1-1) Analyse Graphique	45
1-2) Test de saisonnalité dans Eviews	45
2) Décomposition des séries temporelles dans Eviews	47
3) Modèles de décomposition des séries temporelles	47
3-1) Modèle additif ou modèle multiplicatif	47
3-2) Test de détermination du modèle (Buys-Ballot)	49
4) Le test Dickey-Fuller Augmenté (ADF) et la modélisation VAR sous Eviews	49
4-1) Le test ADF	49
4-2) La modélisation VAR sous Eviews	50
5) Méthodes de désaisonnalisation dans Eviews	51
5-1) Méthode des Moyennes Mobiles	51
5-2) Ajustement par régression avec variables indicatrices	52
5-3) Méthode X-12-ARIMA et X-13-ARIMA-SEATS	53
5-4) Méthode STL (Seasonal-Trend Decomposition using Loess)	54
5-5) Filtrage de Hodrick-Prescott (HP Filter)	54
Exemple d'application	54
Conclusion	59
Chapitre 8 : Fonction d'Autocorrélation (ACF) et Fonction d'Autocorrélation Partielle (PACF)	60
1) Définition et Formules	60
1-1) Fonction d'Autocorrélation (ACF)	60
1-2) Fonction d'Autocorrélation Partielle (PACF)	60
2) Analyse de l'ACF et de la PACF pour les modèles AR, MA et ARMA	61
2-1) Modèle autorégressif (AR(p))	61
2-2) Modèle moyenne mobile (MA(q))	61
2-3) Modèle ARMA(p,q)	61
3) L'autocorrélation et les modèles AR(p), MA(q) et ARMA(p,q) dans Eviews	62
3-1) Générer et Visualiser les Correlogrammes	62
3-2) Ajustement du modèle dans Eviews	62
Application dans Eviews	63
Conclusion	68
Chapitre 9 : Tests de Stabilité	69
Introduction	69
1) Le Test de Chow	69
1-1) Hypothèses du test	69
1-2) Méthodologie	69
1-3) Application sous Eviews	70

2) Tests de Stabilité des Résidus : CUSUM et CUSUMSQ	70
2-1) Test de CUSUM	70
2-2) Test de CUSUMSQ	70
2-3) Application sous Eviews	70
3) Test de Breakpoint de Bai-Perron	70
3-1) Méthodologie	70
3-2) Application sous Eviews	70
Exemple d'application	71
Conclusion	81
Chapitre 10 : La méthodologie de Box-Jenkins	82
Introduction	82
1) Les étapes de la méthodologie Box-Jenkins	82
2) Application sous Eviews	84
Exemple d'application	85
Conclusion	89
Conclusion Générale	90
Références bibliographique	91
Tables des matières	93