

UNIVERSITÉ PARIS-EST CRÉTEIL
UFR DE SCIENCES ÉCONOMIQUES ET DE
GESTION

Master 1 MBFA
Analyse et Modélisation des Risques

Année universitaire 2024–2025

PROJET ÉCONOMÉTRIQUE

PIB Australien

Économétrie appliquée sous SAS

Matière de :
Monsieur Vincent Bouvatier

Réalisé par :

COURTOIS Jamy | MALONGA Dieudonné

Table des matières

Application 1 : Tests de racine unitaire et Outliers	2
Importation des données et création des variables	3
Tests de racine unitaire (efficiente)	5
Graphique et statistiques descriptives	7
Détection des valeurs extrêmes	9
 Application 2 : Modèle économétrique	 11
Génération et analyse des variables explicatives	12
Estimation du modèle	14
Propriétés des résidus	15
Correction et prévisions du modèle	17

Application 1 : Tests de racine unitaire et Outliers

Pour cette première application, nous considérons deux fichiers de données. Ces fichiers proviennent des bases de données de l'OCDE. Le premier fichier porte sur le PIB Australien (Gross Domestic Product - expenditure approach) en niveau (National currency, current prices, annual levels, seasonally adjusted) sur la période 1980Q1–2023Q3. Cette série est exprimée en millions d'unités monétaires (ici le dollar australien), en valeur nominale, et elle est corrigée des variations saisonnières. Ces données sont dans le fichier CSV : GDP.csv. Le second fichier porte sur le déflateur du PIB (Deflator, national base/reference year, seasonally adjusted) Australien sur la même période, 1980Q1–2023Q3. Cette série de déflateur du PIB est exprimée sous forme d'indices (année de base/référence nationale), et elle est également corrigée des variations saisonnières. Ces données sont dans le fichier CSV : Deflator.csv. Nous voulons à l'aide de ces données déterminer l'ordre d'intégration de la série que nous aurons à considérer puis détecter ses éventuelles valeurs extrêmes.

Importation des données et création des variables

Pour réaliser l'importation des données et la création des variables, nous avons effectué plusieurs étapes data sous SAS dont la première a été d'importer les fichiers GDP et Deflator. Nous avons par la suite conservé seulement les données liées à l'Australie. Lors de la dernière étape data, nous avons fusionné les deux bases de données et créés notamment les variables du PIB réel et le logarithme de ce dernier.

Pour obtenir le PIB australien réel, nous faisons :

$$\text{PIB}_{\text{réel}} = \left(\frac{\text{GDP}}{\text{Deflator}} \right) \times 100$$

Nous l'exprimons sous forme logarithmique.

Enfin, nous avons calculé le taux de croissance trimestrielle du PIB de l'Australie comme suit : $\Delta Y_t = \ln(Y_t) - \ln(Y_{t-1})$

Voici l'entête de notre base de données finale :

Obs.	time_var	country	GDP	deflator	Yt	ln_Yt	delta_Yt
1	1980Q1	AUS	136848	20.4843	668064.00	13.4121	.
2	1980Q2	AUS	139964	20.8966	669791.99	13.4147	0.25832
3	1980Q3	AUS	144584	21.4716	673372.01	13.4201	0.53307
4	1980Q4	AUS	151220	22.0778	684940.01	13.4371	1.70333
5	1981Q1	AUS	154248	22.4324	687612.01	13.4410	0.38935

FIGURE 1 – Entête de la base de données finale (source : OCDE)

Nous travaillons sur des séries temporelles. Notre série étant déjà corrigée de ses variations saisonnières, nous allons nous intéresser directement à sa stationnarité. Pour que nous puissions la manipuler, notre série doit être stationnaire. Pour cela, nous allons tester l'ordre d'intégration du logarithme du PIB réel de l'Australie noté $\ln(Y_t)$, en effectuant des tests de racines unitaires. Nous savons qu'une série est intégrée à l'ordre 0 si elle ne nécessite pas de transformation afin d'être stationnaire. Nous allons d'abord commencer par observer le graphique de cette série afin de statuer sur la présomption ou non de stationnarité.

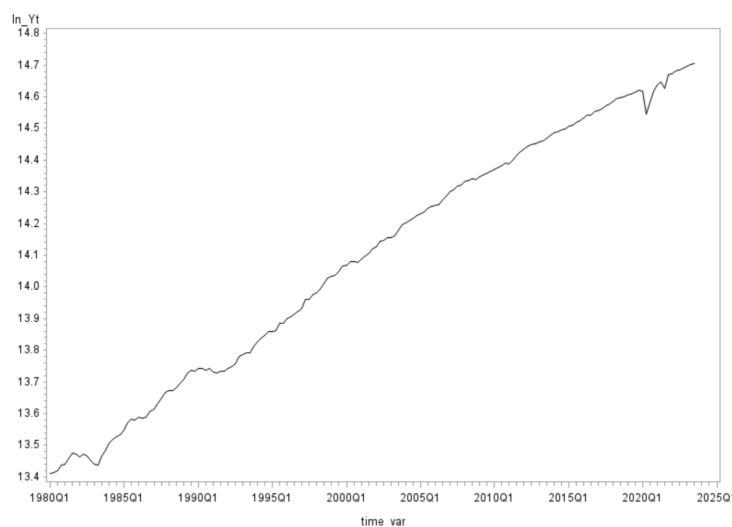


FIGURE 2 – Evolution de la série $\ln(Y_t)$ entre 1980 et 2023

Nous observons une moyenne non constante mais mobile et une variance qui semble, sous réserve d'une échelle moins grande, quant à elle, plutôt constante.

On constate néanmoins, au cours de l'année 2021, une forte volatilité, cela s'explique par le ralentissement de l'économie mondiale lors des nombreux confinements liés à la COVID-19.

Nous avons une présomption de série non-stationnaire mais qui semble suivre une tendance de long terme qui pourrait l'amener à être stationnaire autour de cette dernière.

Nous allons maintenant voir si notre intuition se confirme en effectuant les différents tests de stationnarité adapté.

Tests de racine unitaire (efficente)

Tests de racine unitaire de Dickey-Fuller augmentés							
Type	Retards	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Moyenne zéro	13	0.0882	0.7022	4.09	0.9999	—	—
Moyenne simple	13	-0.5377	0.9235	-2.65	0.0849	13.30	0.0010
Tendance	13	-2.5246	0.9539	-0.76	0.9661	3.67	0.4435

FIGURE 3 – Test ADF sur la série $\ln(Y_t)$

Nous utilisons en premier lieu le test ADF (Augmented Dickey-Fuller) car c'est le premier test de racine unitaire à avoir été développé. Pour ce faire, sur SAS, nous utilisons la proc arima. Nous devons déterminer un hyperparamètre qui est le nombre de retards. Nous avons utilisé le critère de Schertz et il nous indique de sélectionner 13 retards. Nous trouvons cela plus que conservateurs mais c'est le seul critère objectif dont nous disposons afin de sélectionner le nombre de retards.

Voici ce que nous obtenons : Avant de commenter les sorties de ce test, nous allons rappeler les hypothèses de celui-ci.

H_0 : la série a une racine unitaire (elle n'est pas stationnaire)

H_A : la série n'a pas de racine unitaire (elle est stationnaire)

Nous observons que la tendance n'est pas significative alors que la constante l'ait au seuil de 1% car $0,001 < 0,01$. Donc on retient le modèle avec seulement la constante (moyenne simple). On constate que la p-valeur associée est égale à 0,0849 ce qui est supérieur au seuil de significativité de 5%. Nous ne pouvons donc pas rejeter l'hypothèse nulle. Le test nous conduit à statuer sur une éventuelle non-stationnarité de $\ln(Y_t)$. Le test ADF, nous conduit à conclure sur une non-significativité de la tendance.

Néanmoins, nous savons que les tests de première génération de racine unitaire ADF et PP (Phillips-Perron) sont limités lorsque l'on rajoute des éléments déterministes c'est pour cela que nous allons utiliser un test de racine unitaire efficente de deuxième génération : le test ERS (Elliott, Rothenberg et Stock).

Ce test améliore le test ADF. Il permet de tester au sens large si la série est stationnaire par la présence ou non d'une racine unitaire. L'hypothèse nulle et l'hypothèse alternative restent les mêmes.

Sur SAS, nous utilisons une procédure autoreg afin de réaliser le test ERS.

Voici ce que nous obtenons :

Test de racine unitaire Ng-Perron				
Basé sur les données GLS-détendancées spectrales				
Type	MZa	MZt	MSB	MPT
Tendance	-1,7994	-0,8721	0,48464	471,15

FIGURE 4 – Test ERS sur la série $\ln(Y_t)$

Nous avons deux modèles à notre disposition après la réalisation du test, nous choisissons le modèle avec tendance car la tendance visible sur le graphique semble déterministe malgré le test ADF effectué. On voit que la valeur tabulée du test ERS (DF-GLS) est égal à -1,7994 ce qui est inférieur à la statistique de test qui elle vaut

471,15. Nous ne pouvons donc pas rejeter H_0 . Le test ERS, nous indique que la série n'est pas stationnaire.

La série n'est donc pas intégrée d'ordre 0, nous allons effectuer une différenciation et effectuer un test P-P afin de statuer si oui ou non, elle est intégrée d'ordre 1. Nous optons pour ce test plutôt que ADF car il corrige les problèmes liés à l'éventuelle présence d'hétéroscédasticité et d'autocorrélation des erreurs. Les hypothèses restent le même que les tests ADF et ERS.

On pose $\Delta Y_t = \ln(Y_t) - \ln(Y_{t-1})$.

Test de la racine unitaire de Phillips-Perron					
Type	Retards	Rho	Pr < Rho	Tau	Pr < Tau
Moyenne zéro	2	-109,9803	< 0,0001	-8,9851	< 0,0001
Moyenne simp	2	-166,9224	0,0013	-12,8267	< 0,0001
Tendance	2	-167,2600	0,0005	-12,8440	< 0,0001

FIGURE 5 – Test P-P sur la série ΔY_t

Nous choisissons le modèle avec la moyenne simple par intuition graphique. Nous pouvons observer que la p-valeur associée à ce modèle est inférieure à 1%.

Nous rejetons donc H_0 , l'hypothèse de non-stationnarité. La série ΔY_t est stationnaire. La série $\ln(Y_t)$ est donc intégrée d'ordre 1 car il a fallu une différenciation avant de la rendre stationnaire.

Graphique et statistiques descriptives

Après la différenciation, notre série est représentée par le graphique suivant :

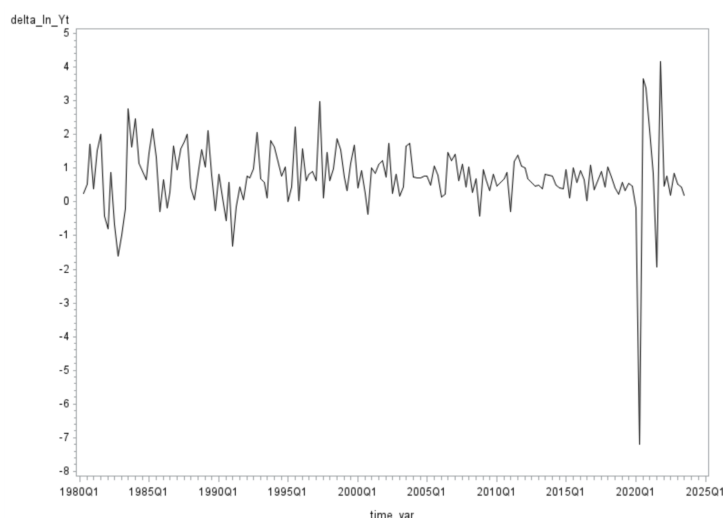


FIGURE 6 – Evolution de la série ΔY_t de 1980 à 2023

La moyenne semble constante autour de 0,8 avec une variance plus au moins comprises entre -2 et 3, hormis lors de forte volatilité au moment des années covid. Cela laissait bien présager que cette série est stationnaire.

Cependant, les valeurs liées aux trimestres des années entre 2020 et 2021 semble bien trop extrêmes et auraient pu nous conduire à statuer sur une série non stationnaire.

Nous allons maintenant nous intéresser aux principales statistiques descriptives de notre série.

Moments			
N	174	Somme des poids	174
Moyenne	0,74208126	Somme des observations	129,12214
Écart-type	1,02265448	Variance	1,04581808
Skewness	-2,9161308	Kurtosis	21,3719824
Somme des carrés non corrigés	276,11155	Somme des carrés corrigés	180,89193
Coeff. Variation	137,79081	Std Error Mean	0,07751967

FIGURE 7 – Statistiques descriptives de la série ΔY_t

Nous disposons de 174 observations. Notre série présente une moyenne d'environ 0.74, cela veut dire qu'en moyenne le taux de croissance du PIB Australien augmente de 0.74% entre 2 trimestres. L'écart-type lui étant associé est d'environ 1.02. Cela correspond à une volatilité plutôt faible. Nous avons calculé indépendamment la médiane sur SAS et elle est d'environ 0.72. La médiane nous indique que 50% des observations que nous possédons sont en-dessous de 0.72% et que 50% sont au-dessus. Le skewness est de -2,35, il n'est pas compris entre -0,5 et 0,5 donc on peut supposer que cette série a un problème d'asymétrie marqué vers la gauche, si tenté qu'elle suive une distribution normale. Le Kurtosis est beaucoup trop élevé (de plus SAS retransche 3 de base donc il est d'environ 24,37).

Une loi normale est censée avoir un Kurtosis de 3, celui-ci est bien trop élevé et vient confirmer la présomption de non-distribution normale qu'on peut émettre (sous réserve que ce dernier ne soit pas biaisé à cause de valeurs extrêmes).

Nous allons maintenant effectuer un test de normalité.

Tests de normalité			
Test	Statistique		p-value
Shapiro-Wilk	$W = 0,795438$	$\Pr < W$	$< 0,0001$
Kolmogorov-Smirnov	$D = 0,13604$	$\Pr > D$	$< 0,0100$
Cramer-von Mises	$W-Sq = 1,115394$	$\Pr > W-Sq$	$< 0,0050$
Anderson-Darling	$A-Sq = 6,25675$	$\Pr > A-Sq$	$< 0,0050$

FIGURE 8 – Test de normalité sur la série ΔY_t

Les hypothèses liées à ces différents tests de normalité sont les suivantes :

H_0 : la série suit une distribution normale

H_A : la série ne suit pas une distribution normale

Bien que nous privilégions le test de Shapiro Wilk (car plus adapté sur des échantillons de faible quantité < 2000), l'ensemble des tests nous conduit à rejeter l'hypothèse nulle.

La distribution de notre série n'est pas normale car les p-valeurs associées au non-rejet de H_0 sont toutes inférieures à 1%.

En revanche, le rejet de cette hypothèse peut être induit par la présence de valeurs extrêmes.

Détection des valeurs extrêmes

Nous allons maintenant identifier ces éventuelles valeurs. Pour ce faire, nous allons effectuer deux approches. La première étant celle du z-score. Cette méthode consiste d'abord à centrer et réduire notre variable. Pour l'effectuer sur SAS, nous utilisons la proc standard à laquelle nous allons fixer la moyenne à 0 et l'écart-type à 1. Par la suite, nous nous assurerons que la moyenne et l'écart-type de la série ΔY_t ont bien été fixé à l'aide de la proc univariate. Nous avons décidé de couvrir 99,75% des observations d'une loi normale. C'est ce qui est le plus conservateur. Nous allons par la suite créer une variable « outlier » sous la forme de la variable binaire qui prendra la valeur de 1 si le z-score associé n'est pas compris entre -3 et 3 et 0 si celui-ci est compris au sein de cette intervalle. Le -3 correspond à la borne inférieure de notre intervalle et le 3 à la borne supérieure de celui-ci.

A l'aide de cette méthode nous obtenons les 2 valeurs extrêmes suivantes :

Obs	time_var	country	GDP	deflator	Yt	ln_Yt	delta_Yt
1	2020Q2	AUS	1875232	90,3397	2075756,01	14,5458	-7,76891
2	2021Q4	AUS	2274872	96,9253	2347206,01	14,6687	3,34220

FIGURE 9 – Détection des valeurs extrêmes sur la série ΔY_t avec la méthode du z-score

Nous avons 2 valeurs extrêmes survenant durant la pandémie mondiale. Comme identifié au moment de l'étape des statistiques descriptives, cette période pouvait prétendre à avoir des valeurs extrêmes car il y a eu une forte volatilité à ce moment. Au deuxième trimestre de 2020, le taux de croissance du PIB Australien a chuté de 7.77% pour ensuite augmenter de 3.34% au dernier trimestre de 2021.

Nous allons maintenant procéder à la détection des valeurs extrêmes avec l'approche de la MAD (Déviation absolue moyenne). Afin de calculer la MAD, nous avons réalisé plusieurs étapes data sur SAS. Tout d'abord, nous avons récupéré la médiane de notre série d'origine. Nous avons ensuite calculé la déviation absolue pour chacune des observations. On a récupéré la médiane de la variable associée à chaque déviation absolue liée à nos observations.

Puis nous avons appliqué la formule de la MAD, et nous obtenons ceci :

Calcul de la MAD	
Médiane de ΔY_t (M)	0,717696
Médiane des $ \Delta Y_t - M $	0,339633
$MAD = 1,4826 \times 0,339633$	0,50354

FIGURE 10 – MAD associée à la série ΔY_t

Nous avons une MAD d'environ 0.503. Afin d'identifier les valeurs extrêmes à l'aide de la MAD, nous appliquons un intervalle de confiance que l'on fixe entre les bornes suivantes :

Médiane de $\Delta Y_t - 2,5 \times \text{la MAD}$ (borne inférieure) et Médiane de $\Delta Y_t + 2,5 \times \text{la MAD}$ (borne supérieure).

Nos observations doivent être comprises au sein de cette intervalle, sinon ce sont des valeurs considérées comme extrêmes par cette méthode d'identification.

Voici les valeurs extrêmes que nous obtenons avec cette méthode :

21 valeurs extrêmes identifiées par la MAD								
Obs	Trim.	ΔY_t	Obs	Trim.	ΔY_t	Obs	Trim.	ΔY_t
1	1980Q1	.	8	1984Q2	-2,035	15	2000Q3	2,087
2	1980Q4	1,703	9	1984Q3	-1,568	16	2003Q1	-1,561
3	1981Q3	-1,633	10	1985Q1	-2,075	17	2007Q4	-1,561
4	1982Q1	-1,964	11	1986Q1	2,226	18	2010Q2	-1,609
5	1982Q3	-2,413	12	1987Q2	2,077	19	2020Q2	-7,769
6	1983Q2	-2,006	13	1993Q1	2,081	20	2020Q3	3,628
7	1983Q4	2,613	14	1997Q2	2,098	21	2021Q4	3,342

FIGURE 11 – Détection des valeurs extrêmes de ΔY_t avec l'approche de la MAD

Nous détectons 21 valeurs extrêmes avec l'aide de cette méthode. Nous avons plus de valeurs extrêmes au moment de la crise sanitaire mondiale (confinement puis relance économique). Nous avons également de nouvelles données extrêmes liées au début des années 1980, cela semble provenir d'une anticipation sur le dollar australien créé en 1966 puis laissé totalement flotter à partir de 1983. Cela a provoqué une baisse importante à moyen terme puis une forte hausse à partir de fin 1983. Une hausse qui marque une poursuite pérenne durant le reste de la décennie 80.

Nous trouvons beaucoup plus de valeurs extrêmes en utilisant la méthode de la MAD. Cette méthode présente de nombreux avantages car elle peut être utilisée peu importe le type de distribution suivie par la série. Elle fonctionne également peu importe la taille de l'échantillon. De plus, contrairement à la méthode du z-score, la MAD n'est pas du tout impactée par les valeurs extrêmes. Le z-score comme on l'utilise avec l'écart-type et la moyenne, il peut être fortement impacté en cas de nombreuses et/ou de fortes valeurs extrêmes.

Notre série ne suivant pas une loi normale, il est encore plus judicieux de privilégier la méthode de la MAD plutôt que celle du z-score. Nous supprimons donc ces valeurs extrêmes afin de poursuivre notre estimation.

Application 2 : Modèle économétrique

Dans cette application, nous avons à considérer le modèle ad hoc suivant :

$$\Delta Y_t = \beta_0 + \beta_1 \Delta Y_{t-1} + \beta_2 \Delta Y_{t-1}^{USA} + u_t$$

avec ΔY_t la différence logarithmique du PIB réel australien entre 2 périodes que nous chercherons à expliquer, ΔY_{t-1}^{USA} représente le taux de croissance réel des USA sur la période précédente. β_0 représente la constante de notre modèle, β_1 est le paramètre associée à la période précédente de notre variable expliquée et u_t est le résidu. Nous devons donc générer nos variables explicatives puis les analyser. Nous estimerons le modèle par la suite, puis nous tenterons de l'améliorer si nos résidus ne respectent pas les caractéristiques classiques (bruit blanc normalement distribué). Nous tenterons d'évaluer, à la fin l'impact des prévisions des taux de croissance réelles des USA, TCEPA, sur notre variable expliquée.

Génération et analyse des variables explicatives

Afin de générer les variables ΔY_{t-1}^{USA} et ΔY_{t-1} , nous avons procédé à diverses étapes data. Nous avons d'abord effectué la même procédure que la première étape de l'application 1 afin d'avoir les données du taux de croissance réelle des USA sur SAS. Par la suite, avons introduit un retard sur la variable ΔY_t en appliquant un lag de 1 sur SAS afin d'obtenir ΔY_{t-1} et on a fait de même avec ΔY_t^{USA} pour avoir ΔY_{t-1}^{USA} . Voici la représentation de ΔY_{t-1}^{USA} :

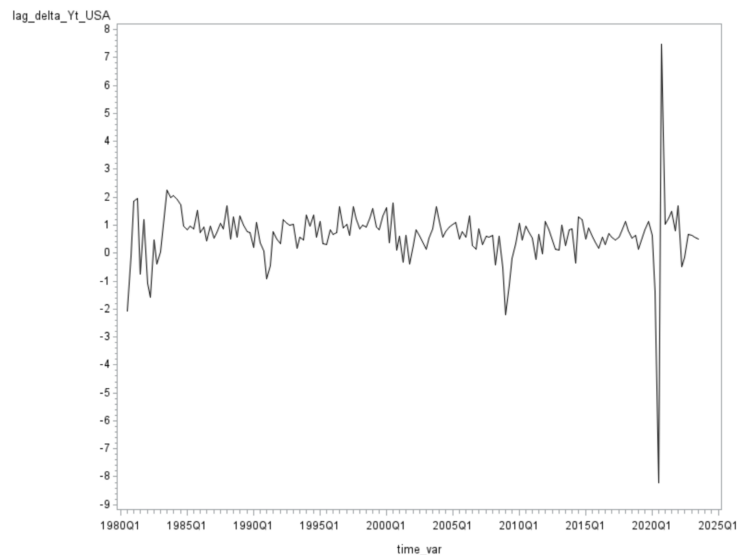


FIGURE 12 – Evolution de ΔY_{t-1}^{USA} de 1980 à 2023

La moyenne semble, tout comme celle sur notre variable que nous voulons expliquer, constante autour de 0,8 avec une variance plus au moins comprise entre -2 et 2 cette fois-ci, hormis lors de forte volatilité au moment des années covid. Nous pouvons déduire que cette série est stationnaire.

Moments			
N	173	Somme des poids	173
Moyenne	0,64027893	Somme des observations	110,768255
Écart-type	1,10537118	Variance	1,22184545
Skewness	-1,905788	Kurtosis	31,554639
Somme des carrés non corrigés	281,079997	Somme des carrés corrigés	210,157417
Coeff. Variation	172,639005	Std Error Mean	0,08403981

FIGURE 13 – Statistiques descriptives de la variable ΔY_{t-1}^{USA}

Nous avons 173 observations et plus 174 car nous avons créé un décalage. La variable a une moyenne d'environ 0,64, cela veut dire qu'en moyenne le taux de croissance du PIB des États-Unis augmente de 0,64% entre 2 trimestres. L'écart-type lui étant associé est d'environ 1,1. Cela correspond à une volatilité plutôt faible. Le skewness est de -1,9, il n'est pas compris entre -0,5 et 0,5 donc on peut supposer que cette série a un problème d'asymétrie marqué vers la gauche, si tant est qu'elle suive une

distribution normale. Le Kurtosis est beaucoup trop élevé (de plus SAS retranche 3 de base donc il est d'environ 35). Tout comme la variable ΔY_t , elle semble ne pas suivre une distribution normale.

Nous allons maintenant voir la potentielle relation qui existe entre ces deux variables afin d'anticiper les résultats de notre estimation.

À partir de ce nuage de points, il ne semble pas exister de relation linéaire entre ces deux variables. Cela va à l'encontre de nos intuitions. Pour être sûr qu'aucune relation linéaire ne se dégage entre ces variables, nous allons déterminer le coefficient de corrélation de ces dernières.

Coefficients de corrélation de Pearson		
Prob $> r $ sous $H_0 : \text{Rho}=0$		
Nombre d'observations		
	delta_Yt	lag_delta_Yt_USA
delta_Yt	1,00000 153	0,09451 0,2468 152
lag_delta_Yt_USA	0,09451 0,2468 152	1,00000 173

FIGURE 14 – Matrice de corrélation entre ΔY_t et ΔY_{t-1}^{USA}

Nous obtenons un coefficient de corrélation inférieur à 0,1. De plus, la p-valeur associée à celui-ci, nous indique qu'il n'est pas significativement différent de 0. Cela vient confirmer, ce qu'on a pu déduire à l'aide du nuage de points ci-dessus. Il ne laisse pas présager une relation linéaire entre ces deux variables.

Estimation du modèle

On a notre modèle qui nous est donné sous la forme de régression ad hoc suivant :

$$\Delta Y_t = \beta_0 + \beta_1 \Delta Y_{t-1} + \beta_2 \Delta Y_{t-1}^{USA}$$

Nous allons l'estimer par la méthode des moindres carrés ordinaires à l'aide de la procédure reg. Nous sommes cependant très pessimistes, quant à la significativité de nos futurs paramètres estimés et au pourcentage de variance expliqué par notre modèle au vu des résultats trouvés précédemment.

Analyse de la variance					
Source	DDL	Somme des carrés	Carré moyen	Valeur F	Pr > F
Modèle	2	0,50774	0,25387	0,68	0,5109
Erreur	149	55,85344	0,37486		
Total corrigé	151	56,36118			

Root MSE	0,61225	R carré	0,0090
Moyenne dép.	0,70204	R carré adj.	-0,0043
Coeff. var.	87,2064		

FIGURE 15 – Modèle par MCO de ΔY_t (1)

Nous nous retrouvons avec une statistique de Fisher extrêmement petite avec une p-valeur associée largement supérieure au seuil de significativité à 5%. Ce qui nous conduit à ne pas rejeter l'hypothèse nulle de non-significativité globale du modèle. Le R^2 est de 0,009, le pouvoir explicatif de notre modèle est extrêmement faible. Notre modèle explique moins de 1% de la variance de notre variable expliquée. De plus, on a un R^2 ajusté négatif, ce qui nous laisse envisager une potentielle erreur de manipulation de nos données ou une mauvaise spécification du modèle.

Résultats estimés des paramètres					
Variable	DDL	Valeur estimée	Erreur type	Valeur du test t	Pr > t
Intercept	1	0,67839	0,07602	8,92	<,0001
lag_delta_Yt	1	-0,01031	0,08681	-0,12	0,9056
lag_delta_Yt_USA	1	0,07591	0,06719	1,13	0,2604

FIGURE 16 – Modèle par MCO de ΔY_t (2)

Nous avons la totalité de nos paramètres estimés qui sont non-significatifs à part la constante. Nos paramètres sont inexploitables. Nous allons quand même voir si cela est dû à une violation d'une des 3 hypothèses classiques sur le terme d'erreurs.

Propriétés des résidus

Test d'hétéroscédasticité :

Pour tester l'hétéroscédasticité de notre modèle, nous effectuons un test de White plutôt qu'un test de Breusch-Pagan car celui-ci couvre un plus grand nombre et de types d'hétéroscédasticité possible.

L'hypothèse nulle de ce test est que le terme d'erreurs est homoscedastique (c'est-à-dire que la variance est la même pour chaque estimation). L'hypothèse alternative est que le terme est hétéroscédastique.

Test de spécification du premier et du deuxième moment		
DDL	Khi 2	Pr > Khi 2
5	2,56	0,7675

FIGURE 17 – Test de White

La p-valeur associée à la statistique du test est supérieure à 10%, nous ne rejetons donc pas H_0 . Le terme d'erreurs est homoscedastique.

Test d'autocorrélation des erreurs :

Pour tester l'autocorrélation du terme d'erreurs de notre modèle, nous effectuons un test de Breusch-Godfrey plutôt qu'un test de Durbin-Watson car ce dernier est beaucoup moins fiable lorsqu'on intègre de la dynamique.

L'hypothèse nulle de ce test est que le terme d'erreurs est non-autocorrélé. L'hypothèse alternative est que le terme d'erreurs est autocorrélé.

Test de corrélation sérielle de Godfrey		
Alternatif	LM	Pr > LM
AR(1)	0,0066	0,9352

FIGURE 18 – Test de Breusch-Godfrey

La p-valeur associée à la statistique du test est supérieure à 10%, nous ne rejetons donc pas H_0 . Le terme d'erreurs est non autocorrélé.

Test de normalité des erreurs :

Pour tester la normalité des erreurs nous utilisons le test de Jarque-Bera.

L'hypothèse nulle de ce test est que le terme d'erreurs est normalement distribué.

L'hypothèse alternative est que le terme d'erreurs n'est pas normalement distribué.

Statistiques diverses			
Statistique	Valeur	Prob	Libellé
Test normal	0,7156	0,6992	Pr > ChiSq

FIGURE 19 – Test de Jarque-Bera

La p-valeur associée à la statistique du test est supérieure à 10%, nous ne rejetons donc pas H_0 . Le terme d'erreurs est normalement distribué.

Correction et prévisions du modèle

La sphéricité du terme d'erreurs étant respectée, tenter de corriger notre modèle à l'aide des estimateurs robustes de Newey et West est inutile. Notre problème vient clairement d'une mauvaise spécification de notre modèle ou bien d'une mauvaise manipulation de nos données. Changer de méthode d'estimation (type moindres carrés généralisés, ridge) ne sert à rien car nous n'avons pas de problèmes liés aux résolutions qu'ils pourraient nous apporter.

Il peut néanmoins exister une relation non linéaire entre nos 2 variables ou bien exister un problème d'endogénéité.

Prévisions :

Voici le calcul des taux de croissance prévisionnels des USA :

Trimestre	Calcul	$\Delta Y_t^{USA} (\%)$
Q4 2024	1,7 / 4	0,425
Q1 2025	1,8 / 4	0,450
Q2 2025	1,9 / 4	0,475
Q3 2025	1,5 / 4	0,375

FIGURE 20 – Taux de croissance prévisionnels des USA (TCEPA)

Ce que l'on aurait pu être tenté de faire en prenant en compte juste nos données est de dire qu'une augmentation de 1% de la valeur de ΔY_{t-1}^{USA} TCEPA entraîne une variation égale au bêta chapeau estimé qui lui correspond sur le taux de croissance de l'Australie. Mais cela semble particulièrement contre intuitif dans notre cas car aucun paramètre n'est interprétable par manque de significativité.