



Rapport d’étude économétrique sur l’investissement des administrations publiques aux Etats-Unis de 1947 à 2022

Enseignante : Mme Ghattassi

Etudiant : Hraouli Samir

Date : 26/03/2023

Master Ingénierie Financière et Modélisation

Table des matières

[I. Introduction 3](#_Toc130756353)

[II. Transformation de la donnée 3](#_Toc130756354)

[III. Représentation des investissements des administrations publiques 3](#_Toc130756355)

[a. Test de stationnarité 3](#_Toc130756356)

[IV. Estimation du modèle ARIMA 4](#_Toc130756357)

[a. ACF et PACF 4](#_Toc130756358)

[b. Test de saisonnalité 5](#_Toc130756359)

[c. Tests de modèles ARIMA 5](#_Toc130756360)

[V. Diagnostic des résidus 6](#_Toc130756361)

[a. Analyse des résidus 7](#_Toc130756362)

[b. Test Box Ljung 7](#_Toc130756363)

[c. Test de Breusch Godfrey 8](#_Toc130756364)

[d. Test de Shapiro 8](#_Toc130756365)

[VI. Prévision de l’investissement des administrations publiques pour 2023 8](#_Toc130756366)

[VII. Conclusion 9](#_Toc130756367)

[VIII. Bibliographie 9](#_Toc130756368)

## Introduction

Ce projet met en lumière les investissements des administrations publiques des Etats-Unis. Les objectifs des investissements publics sont de favoriser la croissance économique en finançant des projets publics tels que des infrastructures publiques. Ce rapport présente une étude économétrique mené sur les investissements des administrations publiques des Etats-Unis depuis 1947 jusqu’à 2022.

## Transformation de la donnée

Dans un premier temps, un travail de remise en forme de la donnée était nécessaire car les données fournies par la bea n’étaient pas exploitable. En effet, celles-ci étaient présentées tels que chaque trimestre correspondent à une colonne. J’ai donc programmé un script sous Python à l’aide des librairies Pandas pour le traitement de Dataframe et Numpy pour générer une séquence de date. Après quelques lignes de code, j’ai pu génerer un dataset, avec des données organisées par ligne, où chaque investissement du gouvernement est assigné à une date, et un trimestre. Le script Python est disponible sur mon github à l’adresse suivante :

[econometrics/projet\_semestre2\_IFIM\_timeseries at main · samhrrl/econometrics (github.com)](https://github.com/samhrrl/econometrics/tree/main/projet_semestre2_IFIM_timeseries)

## Représentation des investissements des administrations publiques

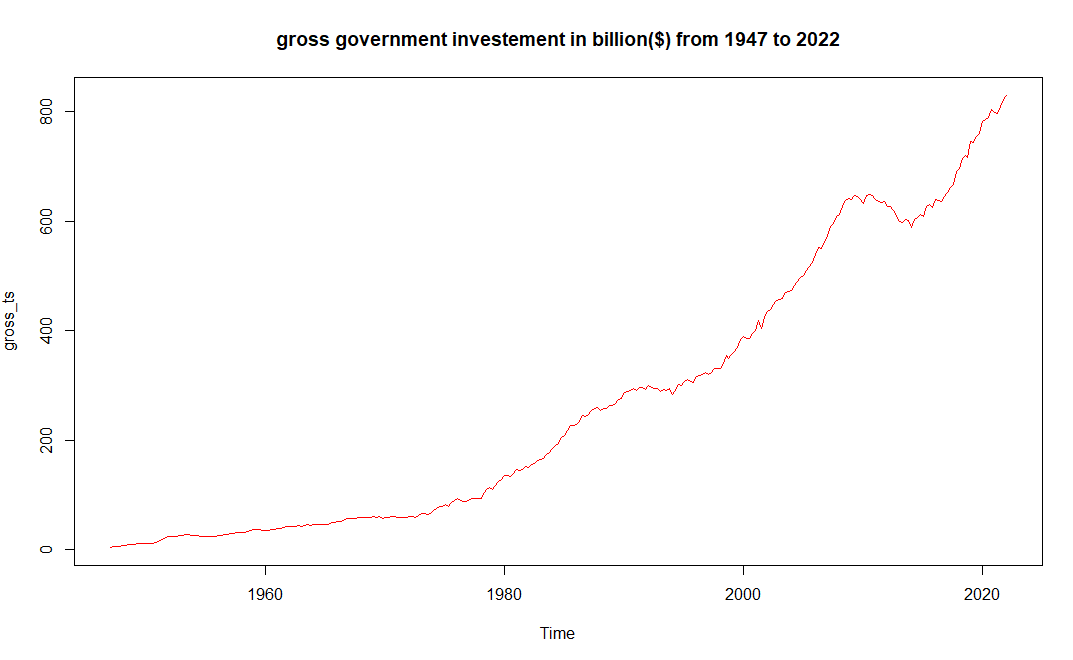
Après avoir transformé la donnée, nous pouvons faire quelques statistiques descriptives et observé comment évolue les investissements des administrations publiques aux Etats-Unis. En effet, on constate que celle-ci est croissante, elle ne cesse d’évoluer chaque année.

Figure 1 Investissements des administrations publiques aux Etats-Unis de 1947 à 2022

Dans les années 80 et 90, les investissements des administrations publiques ont fortement augmenté notamment dans les secteurs de l’éducation, les infrastructures publiques, les soins santé et la sécurité sociale.

En outre, pendant la période covid, les Etats-Unis ont engagés des dépenses massives pour diminuer les effets économiques et sanitaires pour lutter contre cette crise.

Enfin, les élections présidentielles de 2020 aux Etats-Unis ont remis en cause les dépenses et investissements des administrations publiques. En effet, Joe Biden adopte une politque keynésienne, donc il est prêt à relancer l’économie en simulant la croissance du pays en créant de l’emploi et en investissant dans des infrastructures publiques.

### Test de stationnarité

La série présentée montre qu’elle n’est pas stationnaire, nous allons donc la stationnariser en prenant le logarithme de la différence première afin de mener cette étude économétrique.

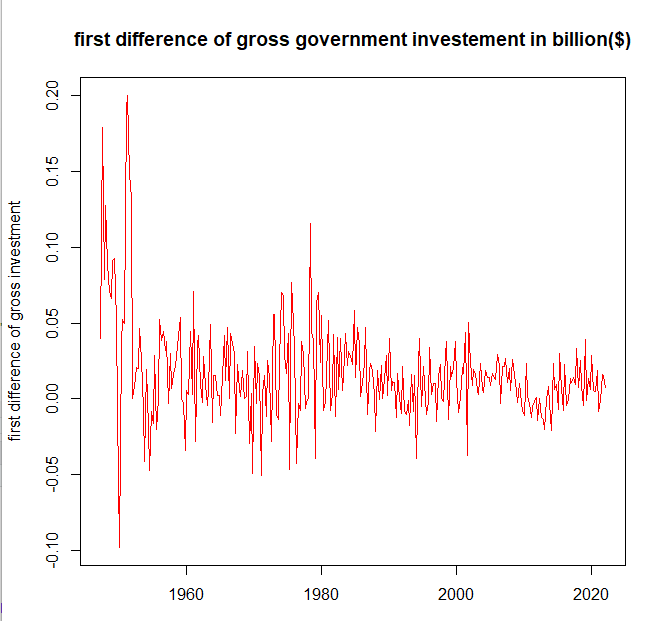


Figure 2 Différence première des investissements des administrations publiques des Etats-Unis

Après une différence première, la série semble stationnaire, nous allons faire un test statistique pour vérifier qu’elle est bien stationnaire.

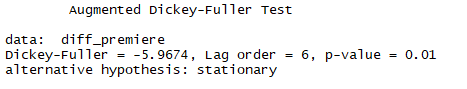


Figure 3 Test de Dickey-Fuller

En réalisant le test de Dickey-Fuller, on constate que la p-value=0.01 < 5%, donc nous pouvons rejeter l’hypothèse nulle qui affirme que la série n’est pas stationnaire.

## Estimation du modèle ARIMA

### ACF et PACF

Nous avons conclus par un test de Dickey-Fuller que les investissements des administrations publiques des Etats-Unis est stationnaire. Nous allons après-en visualiser la fonction d’auto corrélation et leur fonction d’auto corrélation partiel pour estimer un modèle ARIMA.

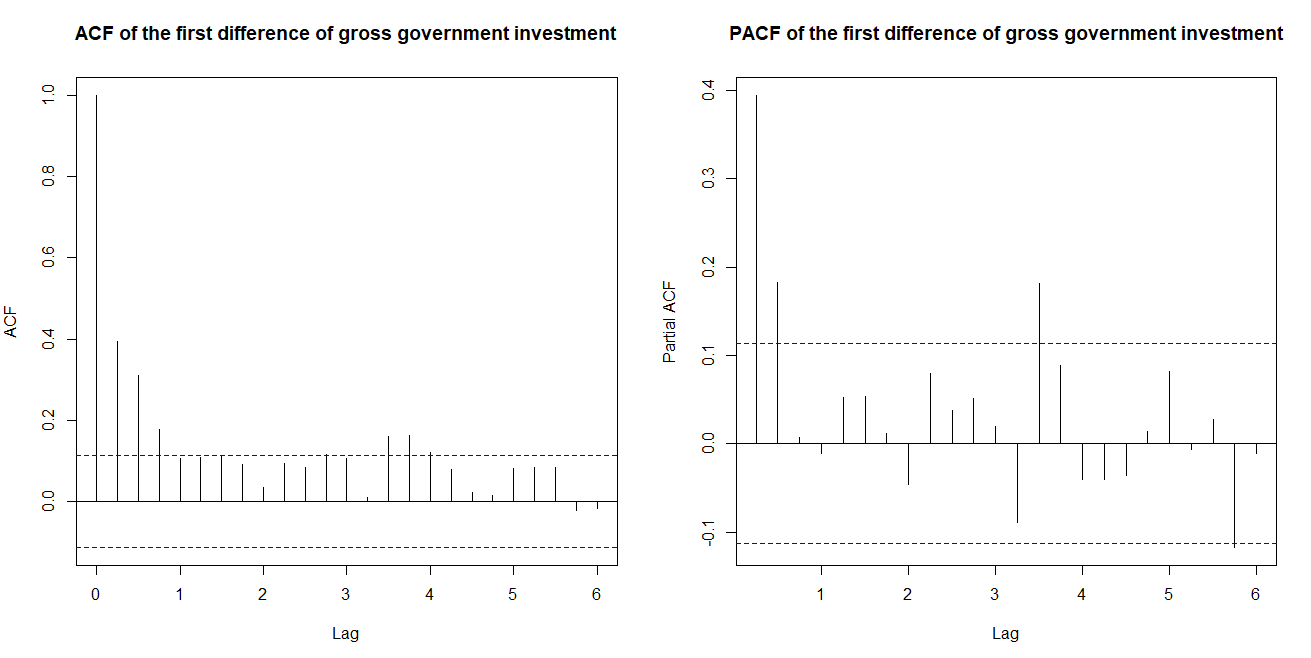


Figure 4 ACF et PACF des investissements des administrations publiques

On constate sur l’ACF, une décroissance forte de nos autos corrélations, on peut donc penser qu’il s’agit d’un modèle autorégressif AR(p). Cependant, on constate un effet de saisonnalité qui se répète régulièrement. Il faudra donc faire un test de saisonnalité pour vérifier s’il y a un aspect saisonnier ou non.

Sur le PACF, on observe un effet sinusoïdale des nos autos corrélations ce qui nous laisse penser qu’il s’agit d’un modèle à moyenne mobile MA(q). De plus, on constate une rupture à l’ordre 2. On peut donc mettre en lumière qu’on aurait un AR(2). Quant à l’ACF, on peut voir une rupture à l’ordre 3 ou 4. On peut donc penser qu’il s’agit d’un MA(3) ou d’un MA(4). Ces spéculations sont établies d’un point de vue conjecturel, nous testerons par la suite, à l’aide de test, quel modèle ARIMA on pourrait modéliser.

### Test de saisonnalité

J’ai réalisé un test statistique qui permet d’évoquer ou non s’il y a un effet saisonnier. J’ai utilisé la librairie (seastests) sur R pour faire un test de saisonnalité. Le test présente deux hypothèses :

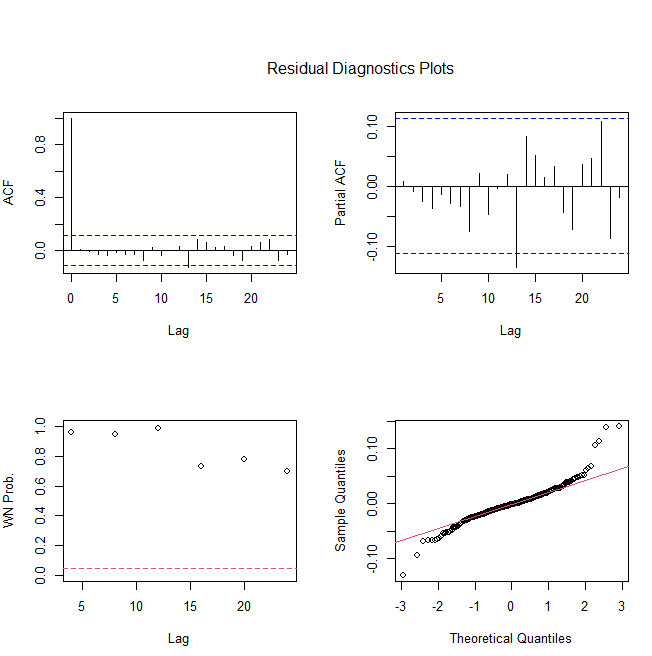
**H0 :** la série présente un effet de saisonnalité, le test retourne un booléen **TRUE**

**He :** la série ne présente pas d’effet de saisonnalité, le test retourne un booléen **FALSE**

Après avoir réalisé le test, j’en ai déduis que la série ne présentait pas d’aspect saisonnier car le test m’a retourné FALSE.

### Tests de modèles ARIMA

Comme la série présente ne présente pas d’aspect saisonnier, j’ai testé plusieurs modèles ARIMA à l’aide de la fonction **estimate** sur R. Pour choisir le modèle, j’ai pris en considération des critères tels que la significativité des t-value, qui devait être supérieur à 1.96 en valeur absolue, la p-value qui devait être significatif au seuil des 10% et le modèle qui minimise le critère d’information AIC. A partir de ces critères, j’ai testé plusieurs modèles, et j’ai constaté lors des premiers tests que la significativité n’était pas toujours correcte, j’ai donc décidé de faire varier mes paramètres p et q en testant des ARIMA(3,0,4), ARIMA(4,0,4), ARIMA(5,0,4), ARIMA(4,0,5) et ARIMA(5,0,3). J’ai conclus que le meilleur modèle était ARIMA (5, 0,3), car mes résultats étaient significatifs sauf pour l’AR(4) où ma t-value =-0.775 < 1.96 et AR(5) = -1.461<1.96, Seul, le MA(3) présente un résultat non significatif sur sa p-value, mais j’ai décidé d’accepter le modèle car les autres valeurs étaient acceptables et que ce modèle minimisait également le critère d’information AIC= -1239.748.



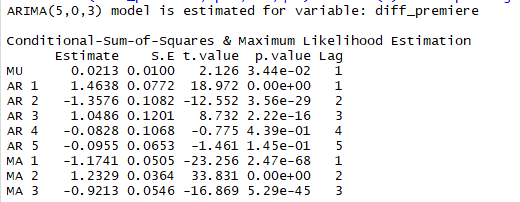


Figure 5 Estimation d’un modèle ARIMA(5,0,3)

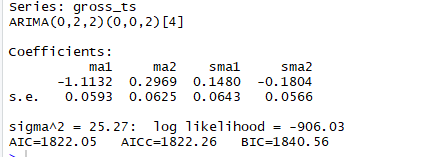


Figure 6 Résidus du modèle ARIMA(5,0,3)

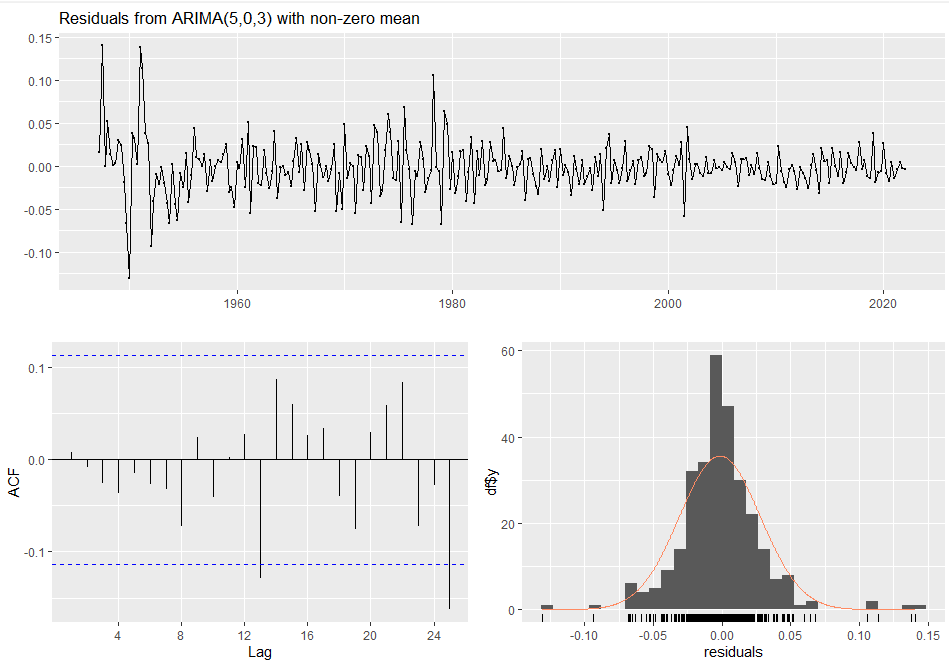
Figure 7 Estimation d'un modèle ARIMA automatique sous R

Enfin, j’ai également interrogé Rstudio pour qu’il me propose un modèle automatique ARIMA. Par conséquent, le modèle proposé est un ARIMA (0,2, 2). J’ai rejeté ce modèle car celui-ci ne contient pas de composante AR. Le modèle fourni par R est un MA pur qui ne contient pas de données observables mais uniquement un ensemble de chocs. De plus, le modèle proposé par R utilise la différence seconde, alors que la série est stationnaire après une différence première. Donc, on perd de l’information.

## Diagnostic des résidus

Nous allons réaliser un diagnostic des résidus afin de s’assurer que le modèle ARIMA choisi est approprié pour faire des prévisions futures.

### Analyse des résidus



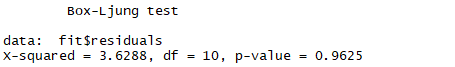
D’après le graphique, on constate que les résidus se concentrent autour de la moyenne. De plus, on observe un pic conséquent dans les années 45, En effet, pendant la période post seconde guerre mondiale, les Etats-Unis ont massivement dépensé pour leur défense mais aussi pour le soutien apporté à leurs alliés. De plus, après la guerre, le gouvernement a investi dans la reconstruction de l’Europe dans le cadre du plan Marshall et a commencé à investir dans des projets plus complexes tels que la construction d’infrastructure, l’innovation et la recherche scientifique ainsi que l’éducation ; Ceci peut expliquer pourquoi la volatilité est haute pendant cette période.

### Test Box Ljung

Nous allons ensuite réaliser un **test d’indépendance des résidus**. Ce test est constitué de deux hypothèses :

**H0 :** la corrélation entre les résidus est très faible (indépendance)

**He :** la corrélation entre les résidus est significativement différente de 0(il n’y a pas indépendance)



J’en conclus, avec le test de Box Ljung que la p-value=0.9625, ce résultat étant supérieur à 10%, nous pouvons donc rejeter l’hypothèse He que la corrélation entre les résidus est significativement différente de 0. Donc, les résidus sont indépendants.

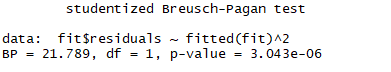
### Test de Breusch Godfrey

Nous allons ensuite faire un test d’homoscédasticité des résidus afin de vérifier que nos résidus sont homoscédastiste, c'est-à-dire que leur variance ne varie pas dans le temps. Nous utiliserons le test de Breusch Godfrey sur R. Ce teste présente deux hypothèses :

**H0 :** les résidus présente un cas d’homoscédasticité, la variance est donc constante dans le temps

**He :** les résidus présente un cas d’hétéroscédasticité, la variance varie dans le temps.

Après avoir réalisé le test, nous obtenons les résultats suivants

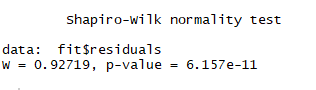


Nous constatons que la p-value=3.04x10E(-6), est inférieur au seuil des 10%. Donc, je rejette l’hypothèse nulle h0. Ainsi, la série n’est pas un cas d’homoscédasticité, la variance varie dans le temps, il s’agit d’un cas d’hétéroscédasticité.

Celle-ci est rejeté en raison de la haute volatilité de nos résidus notamment le pic conséquent qu’on retrouve pendant la période après guerre.

### Test de Shapiro

Nous finirons par un test de normalité des résidus pour savoir si les résidus suivent ou non une distribution normale avec un test de Shapiro. On constate après les résultats du test, que la p-value est faible : 6.15x10E (-11), qui est inférieur au seuil des 10%, j’en conclus que les résidus ne suivent pas une distribution normale.



## Prévision de l’investissement des administrations publiques pour 2023

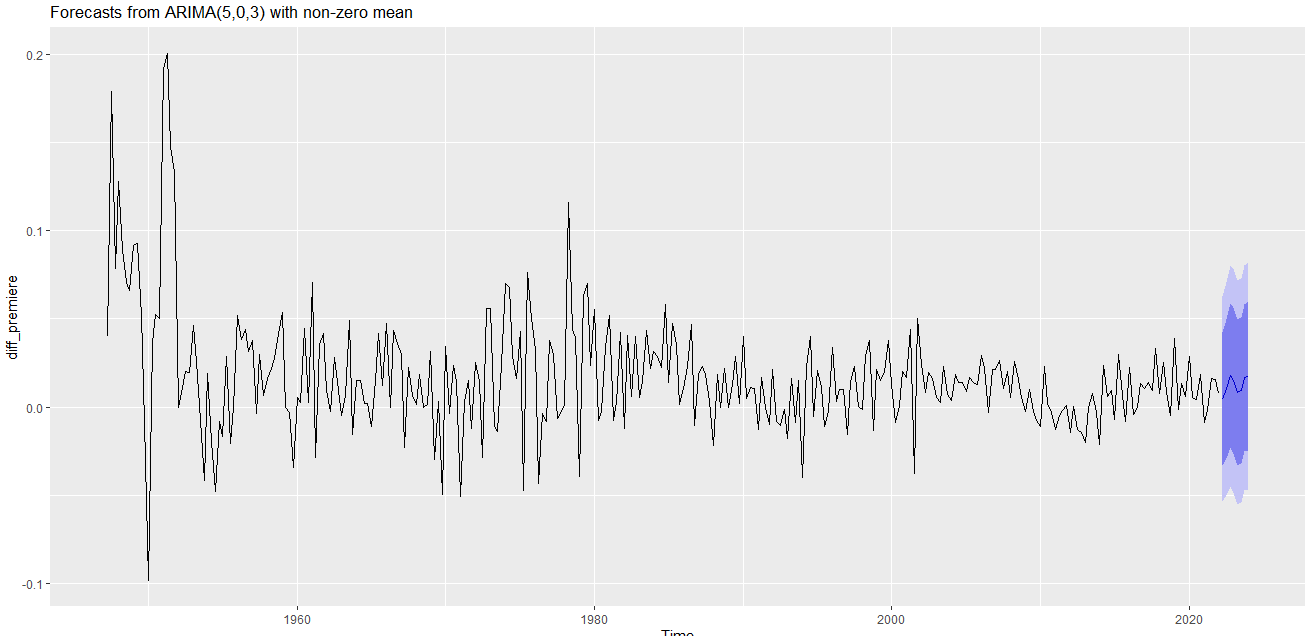
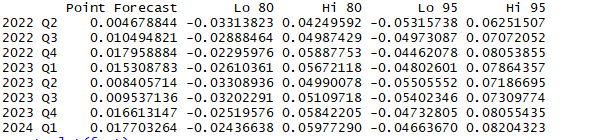


Figure Prévision des investissements des administrations publiques pour l'année 2023

Nous pouvons alors faire des prévisions pour 2023 en utilisant la librairie forecast. On constate que l’investissement des administrations publiques semble augmenter en début d’année et diminuer par la suite avant de recroitre au trimestre 1 de 2024.

Nous pouvons considérer cette augmentation des investissements car les Etats-Unis sont fortement impliqué dans le conflit de guerre en Ukraine en leur versant des aides financières et matérielles qui s’élèvent à 50 milliard de dollars.



## Conclusion

Nous avons mené cette étude économétrique en estimant et en faisant une prévision des futures valeurs de l’investissement des administrations publiques aux Etats-Unis avec le meilleur modèle ARIMA en tenant compte de la significativité de nos résultats ainsi que le modèle minimisant le critère d’information AIC. Toutefois, le modèle ARIMA ne serait pas le plus adapté en raison de la haute volatilité de notre modèle, où le test d’homoscédasticité est rejeté. Il serait donc préférable d’orienter cette étude en utilisant un modèle GARCH ou ARCH qui est un modèle reposant sur la haute volatilité.

## Bibliographie

[États-Unis – Décision de la FED sur les taux d'intérêt (investing.com)](https://fr.investing.com/economic-calendar/interest-rate-decision-168?__cf_chl_tk=616KgdVUBmRqqfxFYHSS6jgt4MeGgUIYpL7FIsoUsyM-1679848676-0-gaNycGzNGFA)

[8.7 ARIMA modelling in R | Forecasting: Principles and Practice (2nd ed) (otexts.com)](https://otexts.com/fpp2/arima-r.html)

[r - Plot forecast and actual values - Stack Overflow](https://stackoverflow.com/questions/59136444/plot-forecast-and-actual-values)

[Nouveau sursaut de l’inflation en janvier aux États-Unis | Le Devoir](https://www.ledevoir.com/economie/781653/nouveau-sursaut-de-l-inflation-en-janvier-aux-etats-unis)

[Le keynésianisme impérialiste de Joe Biden - CONTRETEMPS](https://www.contretemps.eu/keynesianisme-imperialiste-biden/)

[gov\_glance-2013-23-fr.pdf (oecd-ilibrary.org)](https://www.oecd-ilibrary.org/docserver/gov_glance-2013-23-fr.pdf?expires=1679860326&id=id&accname=guest&checksum=AE43B33FA183E483A83EBABBE1917BC8)