

# Analyse des déterminants macroéconomique du déficit public en France

KINDBEITER Arnaud, HOBALLAH Rayan, ZELLER Emile

2024-12-08

# Introduction

Le déficit public en France est influencé par divers facteurs macroéconomiques. L'objectif de cette analyse est de comprendre **quels sont les principaux déterminants qui influencent le déficit public au cours des 20 dernières années**. Les variables étudiées incluent le taux de chômage, le taux d'épargne des ménages, le taux d'inflation, le taux de change effectif, les recettes fiscales nettes et le PIB.

## Description des variables

Les données utilisées pour cette analyse proviennent d'Eurostat, accessibles via ce lien. Ces données couvrent plusieurs indicateurs économiques essentiels pour comprendre l'impact de la crise sanitaire sur les paramètres macroéconomiques.

- **Government.surplus.or.deficit** : Déficit/excédent public
- **NetTaxReceipts** : Total des recettes provenant des impôts et des cotisations sociales (source)
- **Unemployment.Rate** : Taux de chômage (trimestriel, agrégé en annuel pour notre étude)
- **effective.exchange.rate** : Taux de change effectifs des pays industrialisés
- **Household.saving.rate** : Taux d'épargne brut des ménages
- **InflationRate** : Indice des prix à la consommation harmonisé
- **GDP** : PIB et principales composantes (production, dépenses et revenus)
- **year** : Les différentes années étudiées

## Préparation des données

On créer un data frame avec les données qui seront utiliser dans notre modèle.

```
data <- data.frame(
  `effective.exchange.rate` = c(100.23875, 100.23489, 100.22141, 100.19603, 100.11817, 100.04099, 100.00000),
  `Household.saving.rate` = c(14.34, 14.45, 13.71, 14.05, 14.41, 14.38, 15.73, 15.43, 15.05, 15.19, 13.71),
  `GDP` = c(1, 2.9, 1.9, 2.7, 2.5, 0.4, -2.8, 2, 2.4, 0.2, 0.8, 1, 1.1, 0.9, 2.1, 1.6, 2, -7.4, 6.9, 2.4),
  `Unemployment.Rate` = c(8.500, 8.875, 8.875, 8.850, 8.025, 7.425, 9.125, 9.275, 9.225, 9.775, 10.275, 10.275, 10.275, 10.275, 10.275, 10.275, 10.275, 10.275, 10.275, 10.275),
  `Government.surplus.or.deficit` = c(-4.1, -3.6, -3.5, -2.7, -3.0, -3.5, -7.4, -7.2, -5.3, -5.2, -4.9, -4.9, -4.9, -4.9, -4.9, -4.9, -4.9, -4.9, -4.9, -4.9),
  `NetTaxReceipts` = c(44.3, 44.4, 44.7, 45.1, 44.6, 44.5, 44.1, 44.2, 45.3, 46.5, 47.4, 47.6, 47.7, 47.7, 47.7, 47.7, 47.7, 47.7, 47.7, 47.7),
  `InflationRate` = c(2.18, 2.35, 1.9, 1.89, 1.6, 3.17, 0.11, 1.7, 2.3, 2.23, 0.99, 0.62, 0.08, 0.31, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0),
  `year` = c(2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022)
)
```

On visualise la base de données afin de vérifier qu'il ne manque pas de valeurs...

```
mon_data_frame
```

```
##      year effective.exchange.rate Household.saving.rate InflationRate
## 1  2003                100.23875                14.34         2.18
## 2  2004                100.23489                14.45         2.35
## 3  2005                100.22141                13.71         1.90
## 4  2006                100.19603                14.05         1.89
## 5  2007                100.11817                14.41         1.60
## 6  2008                100.04099                14.38         3.17
## 7  2009                100.00065                15.73         0.11
## 8  2010                100.00136                15.43         1.70
## 9  2011                100.00070                15.05         2.30
## 10 2012                 99.99810                15.19         2.23
## 11 2013                 99.99955                13.87         0.99
## 12 2014                100.00000                14.15         0.62
## 13 2015                100.00000                13.61         0.08
## 14 2016                100.00000                13.41         0.31
## 15 2017                100.00000                13.67         1.17
## 16 2018                100.00000                13.54         2.10
## 17 2019                100.00000                14.24         1.31
## 18 2020                100.00000                20.03         0.54
## 19 2021                100.00000                18.67         2.08
## 20 2022                100.00000                16.46         5.90
## 21 2023                100.00000                16.52         5.70
##      Unemployment.Rate NetTaxReceipts  GDP Government.surplus.or.deficit
## 1                   8.500           44.3   1.0                   -4.1
## 2                   8.875           44.4   2.9                   -3.6
## 3                   8.875           44.7   1.9                   -3.5
## 4                   8.850           45.1   2.7                   -2.7
## 5                   8.025           44.6   2.5                   -3.0
## 6                   7.425           44.5   0.4                   -3.5
## 7                   9.125           44.1 -2.8                   -7.4
## 8                   9.275           44.2   2.0                   -7.2
## 9                   9.225           45.3   2.4                   -5.3
## 10                  9.775           46.5   0.2                   -5.2
## 11                 10.275           47.4   0.8                   -4.9
## 12                 10.250           47.6   1.0                   -4.6
## 13                 10.375           47.7   1.1                   -3.9
## 14                 10.075           47.7   0.9                   -3.8
## 15                  9.425           48.5   2.1                   -3.4
## 16                  9.025           48.3   1.6                   -2.3
## 17                  8.425           47.3   2.0                   -2.4
## 18                  8.025           47.4 -7.4                   -8.9
## 19                  7.900           46.9   6.9                   -6.6
## 20                  7.350           47.6   2.6                   -4.7
## 21                  7.350           45.6   0.9                   -5.5
```

## Analyse préliminaire

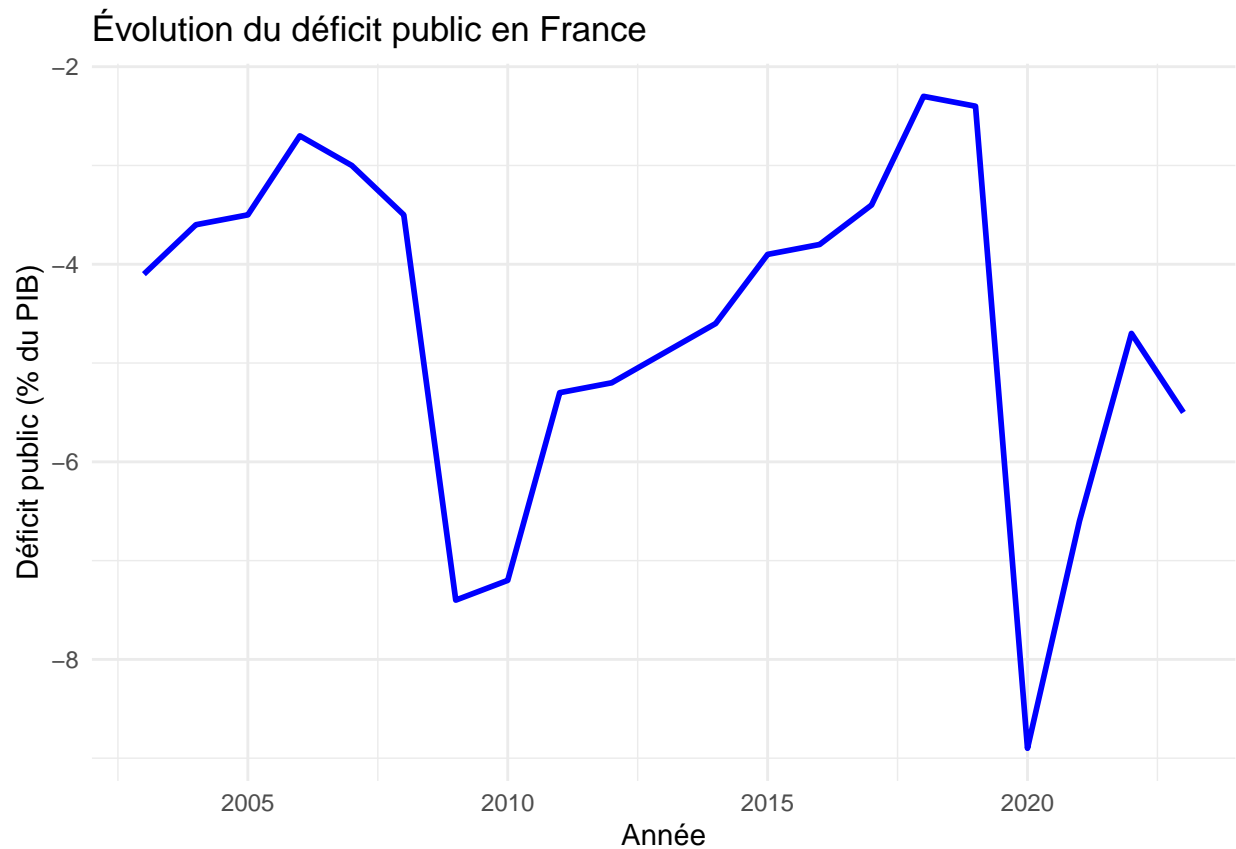
Les statistiques descriptives nous offrent un aperçu global des caractéristiques des variables de notre modèle. Elles permettent d'examiner la répartition des données, les tendances générales, ainsi que les éventuelles valeurs aberrantes ou incohérentes. Ces premières étapes sont importantes pour comprendre le comportement des variables et préparer les analyses plus approfondies qui suivront.

```
summary(mon_data_frame)
```

```
##      year      effective.exchange.rate Household.saving.rate InflationRate
## Min.   :2003      Min.   :100.0           Min.   :13.41           Min.   :0.080
## 1st Qu.:2008      1st Qu.:100.0           1st Qu.:13.87           1st Qu.:0.990
## Median :2013      Median :100.0           Median :14.38           Median :1.890
## Mean   :2013      Mean   :100.1           Mean   :15.00           Mean   :1.916
## 3rd Qu.:2018      3rd Qu.:100.0           3rd Qu.:15.43           3rd Qu.:2.230
## Max.   :2023      Max.   :100.2           Max.   :20.03           Max.   :5.900
## Unemployment.Rate NetTaxReceipts      GDP
## Min.   : 7.350      Min.   :44.10      Min.   : -7.400
## 1st Qu.: 8.025      1st Qu.:44.60      1st Qu.: 0.900
## Median : 8.875      Median :46.50      Median : 1.600
## Mean   : 8.877      Mean   :46.18      Mean   : 1.224
## 3rd Qu.: 9.425      3rd Qu.:47.60      3rd Qu.: 2.400
## Max.   :10.375      Max.   :48.50      Max.   : 6.900
## Government.surplus.or.deficit
## Min.   : -8.900
## 1st Qu.: -5.300
## Median : -4.100
## Mean   : -4.595
## 3rd Qu.: -3.500
## Max.   : -2.300
```

## Visualisation des tendances au fil des années

```
ggplot(mon_data_frame, aes(x = year, y = Government.surplus.or.deficit)) +
  geom_line(color = "blue", size = 1) +
  labs(title = "Évolution du déficit public en France",
       x = "Année",
       y = "Déficit public (% du PIB)") +
  theme_minimal()
```



Le graphique montre des fluctuations marquées du déficit public en France, avec un creux notable en 2009, dû à la crise financière mondiale. Après un redressement progressif jusqu'en 2019, le déficit s'aggrave brutalement en 2020 en raison de la crise du COVID-19, atteignant près de 10 % du PIB. Bien qu'en légère réduction en 2021, le déficit reste supérieur à son niveau d'avant-crise. Cela illustre l'impact des crises économiques sur les finances publiques françaises.

#### Matrice de corrélation

```
cor_matrix <- cor(data[, c("Unemployment.Rate", "Household.saving.rate", "effective.exchange.rate", "Ne
```



La matrice de corrélation montre des relations linéaires entre les variables macroéconomiques. Une corrélation négative notable est observée entre le taux de chômage et le taux d'inflation (-0,69), suggérant une relation de type courbe de Phillips. Le taux d'épargne des ménages est également négativement corrélé au chômage (-0,53), ce qui peut refléter une réduction de l'épargne en période de chômage élevé. Enfin, une corrélation négative modérée entre le taux de change effectif et les recettes fiscales nettes (-0,59) indique un impact potentiel du taux de change sur les flux fiscaux. Cependant, ces corrélations ne traduisent pas de causalité et nécessitent des analyses économétriques plus approfondies.

## Estimation du modèle de régression

Pour estimer l'impact des déterminants macroéconomiques sur le déficit public, nous avons utilisé un modèle de régression linéaire multiple. Ce modèle évalue l'influence des variables suivantes : le taux de chômage (Unemployment.Rate), le taux d'épargne des ménages (Household.saving.rate), le taux de change effectif (effective.exchange.rate), les recettes fiscales nettes (NetTaxReceipts), le PIB (GDP), et le taux d'inflation (InflationRate) sur le solde budgétaire (Government.surplus.or.deficit).

```
model <- lm(Government.surplus.or.deficit ~ Unemployment.Rate + Household.saving.rate + effective.exchan

##
## Call:
## lm(formula = Government.surplus.or.deficit ~ Unemployment.Rate +
##     Household.saving.rate + effective.exchange.rate + NetTaxReceipts +
##     GDP + InflationRate, data = data)
##
## Residuals:
```

```

##      Min      1Q   Median      3Q      Max
## -0.65638 -0.33854  0.05501  0.28288  0.83580
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -530.06680   165.92983   -3.195  0.00649 **
## Unemployment.Rate    -1.10398    0.20304   -5.437  8.75e-05 ***
## Household.saving.rate   -0.99688    0.08879  -11.228  2.19e-08 ***
## effective.exchange.rate    5.26219    1.62633    3.236  0.00598 **
## NetTaxReceipts     0.51723    0.09013    5.739  5.13e-05 ***
## GDP                0.10324    0.04732    2.182  0.04666 *
## InflationRate      -0.14142    0.10791   -1.310  0.21112
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.481 on 14 degrees of freedom
## Multiple R-squared:  0.9472, Adjusted R-squared:  0.9246
## F-statistic: 41.89 on 6 and 14 DF,  p-value: 3.729e-08

```

## Analyse des résultats du modèle de régression

Le **coefficient de l'Intercept** est **négatif et significatif** (p-value = 0.00649). Cela signifie que lorsque toutes les variables explicatives (chômage, épargne des ménages, taux de change, recettes fiscales nettes, PIB et taux d'inflation) sont égales à 0, le modèle prédit un déficit du gouvernement de 530 unités. Cette valeur n'a pas nécessairement de signification économique claire en dehors de l'échelle des variables utilisées.

Le **coefficient du taux de chômage** est **négatif et significatif** (p-value = 8.75e-05), ce qui suggère qu'une augmentation du taux de chômage de 1 point de pourcentage entraîne une augmentation du déficit public de 1.10 unités, ceteris paribus.

Le **coefficient du taux d'épargne des ménages** est également **négatif et fortement significatif** (p-value = 2.19e-08). Cela implique qu'une augmentation du taux d'épargne des ménages de 1 point de pourcentage est associée à une augmentation du déficit public de 0.997 unités, toutes choses égales par ailleurs.

Le **coefficient du taux de change effectif** est **positif et significatif** (p-value = 0.00598). Cela signifie qu'une appréciation de la monnaie de 1 point entraîne une réduction du déficit public (ou une augmentation du surplus) de 5.26 unités.

Le **coefficient des recettes fiscales nettes** est **positif et significatif** (p-value = 5.13e-05). Une augmentation des recettes fiscales nettes de 1 unité est associée à une réduction du déficit public (ou une augmentation du surplus) de 0.52 unités.

Le **coefficient du PIB** est **positif et significatif** (p-value = 0.04666). Cela signifie que, toutes choses égales par ailleurs, une augmentation de 1 unité du PIB entraîne une augmentation de 0.10324 unités dans le surplus gouvernemental.

Le **coefficient du taux d'inflation** est **négatif et non significatif** (p-value = 0.21112). Cela indique que le taux d'inflation n'a pas d'impact statistiquement significatif sur le surplus ou déficit gouvernemental dans ce modèle.

## R<sup>2</sup> ajusté et Test F

Le **R<sup>2</sup> ajusté** du modèle est de **0.9246**, indiquant que **92.46 %** de la variabilité du déficit public est expliquée par les variables incluses dans le modèle, après ajustement pour le nombre de variables.

Le **test F** du modèle est de **41.89** avec une **p-value de 3.729e-08**, ce qui est extrêmement faible. La faible p-value permet de **rejeter l'hypothèse nulle**, montrant ainsi que les **variables sélectionnées ont un impact globalement significatif** sur le déficit public. Le modèle dans son ensemble est donc **statistiquement pertinent**.

## Tests de diagnostic

### Test de multicolinéarité

```
vif_values <- vif(model)
```

##	Unemployment.Rate	Household.saving.rate	effective.exchange.rate
##	3.298781	2.008383	1.860130
##	NetTaxReceipts	GDP	InflationRate
##	1.689468	1.330532	2.367782

Si ( $VIF < 5$ ) Pas de multicolinéarité importante détectée. Si ( $VIF > 5$ ) Multicolinéarité modérée Si ( $VIF > 10$ ) Multicolinéarité préoccupante

Dans notre cas, toutes les valeurs sont bien inférieures à 5. On ne fait donc pas face à une multicolinéarité importante.

### Test d'hétéroscédasticité

Pour évaluer la présence d'hétéroscédasticité dans le modèle, nous utilisons le test de Breusch-Pagan. Ce test repose sur les hypothèses suivantes :

- H : Les résidus sont homoscedastiques (variance constante).
- H : Les résidus présentent une hétéroscédasticité (variance non constante).

Une p-valeur associée au test est interprétée comme suit :

Si  $p > 0,05$  : Nous ne rejetons pas H . Les résidus sont considérés comme homoscedastiques. Si  $p \leq 0,05$  : Nous rejetons H . Les résidus présentent une hétéroscédasticité significative, ce qui peut indiquer une violation des hypothèses du modèle linéaire.

```
bptest(model)
```

```
##
## studentized Breusch-Pagan test
##
## data: model
## BP = 2.8334, df = 6, p-value = 0.8294
```

La p-value obtenue (0.8294) est bien supérieure au seuil de signification de 5% (0.05). Nous pouvons donc conclure que les résidus du modèle sont homoscedastiques, c'est-à-dire que la variance des erreurs est constante et ne présente pas de problèmes d'hétéroscédasticité.

## Reset test de Ramsey

Pour tester si le modèle est correctement spécifié et qu'il n'y a pas de variables omises importantes, nous utilisons le RESET Test de Ramsey, qui permet de vérifier la validité de la spécification du modèle.

```
resettest(model)
```

```
##  
## RESET test  
##  
## data: model  
## RESET = 0.033759, df1 = 2, df2 = 12, p-value = 0.9669
```

On obtient une statistique de 0.033759 avec une p-value de 0.9669. Cette p-value étant bien supérieure au seuil de 0.05, nous ne rejetons pas l'hypothèse nulle, ce qui suggère que le modèle est correctement spécifié et qu'il n'y a pas de variables omises importantes.

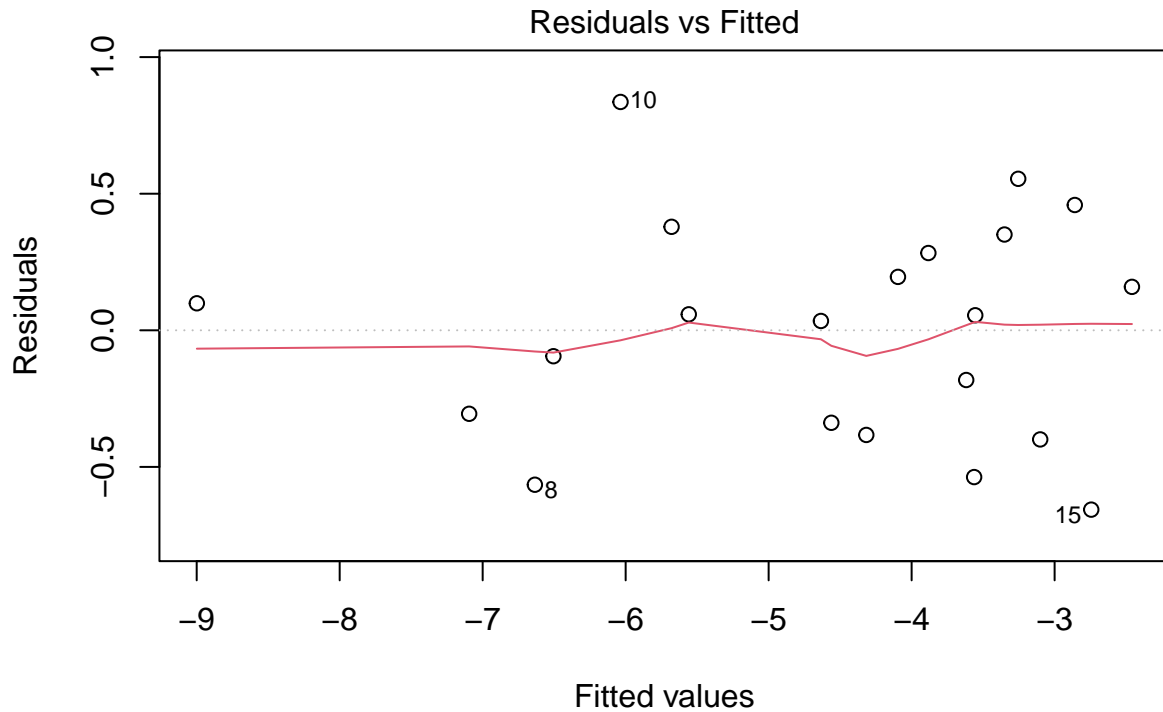
## Analyse des résidus

L'analyse des résidus est une étape cruciale pour évaluer la qualité de l'ajustement de notre modèle. Les résidus sont les différences entre les valeurs observées et les valeurs ajustées par le modèle, et leur comportement peut nous indiquer si certaines hypothèses du modèle sont violées.

**Résidus vs valeurs ajustées** Un premier diagnostic visuel consiste à examiner la relation entre les résidus et les valeurs ajustées.

```
plot(model, which = 1)
```



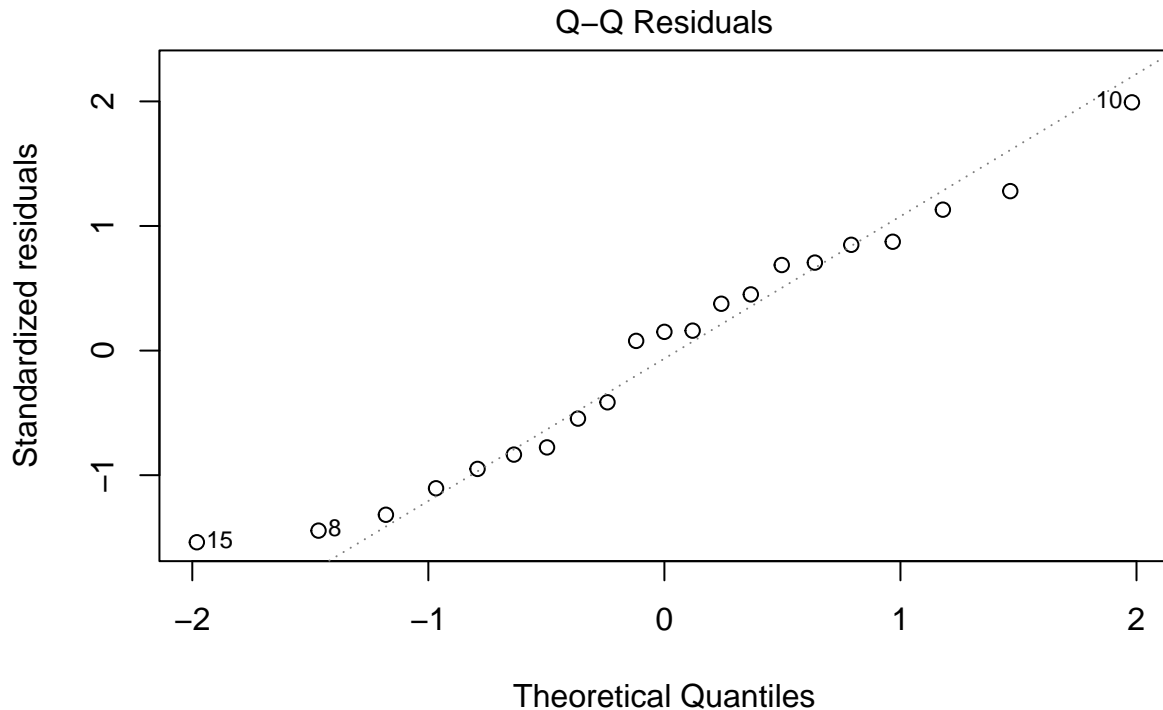


`lm(Government.surplus.or.deficit ~ Unemployment.Rate + Household.saving.rat ...`

Vu que les résidus sont répartis de manière aléatoire autour de zéro, cela nous suggère que le modèle est bien spécifié.

**QQ-plot des résidus** Représentons le QQ-plot afin d'évaluer si les résidus suivent une distribution normale.

```
plot(model, which = 2)
```



$\text{lm}(\text{Government.surplus.or.deficit} \sim \text{Unemployment.Rate} + \text{Household.saving.rat} \dots)$

Vu que les points du graphique suivent une droite diagonale, cela suggère que les résidus sont normalement distribués.

```
shapiro.test(residuals(model))
```

#### Test de normalité des résidus

```
##
## Shapiro-Wilk normality test
##
## data: residuals(model)
## W = 0.97551, p-value = 0.849
```

Le test de Shapiro-Wilk renvoie une statistique  $W=0.97551$  et une p-value de 0.849. Comme la p-value est largement supérieure au seuil de signification de 0.05, nous ne pouvons pas rejeter l'hypothèse nulle de normalité des résidus. Cela indique que les résidus sont normalement distribués, ce qui renforce la validité de notre modèle.

**Autocorrélation des résidus** Pour évaluer la présence d'autocorrélation des résidus dans le modèle, nous utilisons le statistique de Durbin-Watson. Ce test repose sur les hypothèses suivantes :

- $H$  : Absence d'autocorrélation des résidus.

- $H$  : Autocorrélation positive des résidus.

La statistique de Durbin-Watson (DW) varie entre 0 et 4, et son interprétation est la suivante :

- $DW = 2$  : Absence d'autocorrélation.
- $DW < 2$  : Présence d'une autocorrélation positive.
- $DW > 2$  : Présence d'une autocorrélation négative.

Nous nous intéressons également à la valeur  $p$  associée au test :

Si  $p > 0,05$  : L'autocorrélation des résidus n'est pas significative au seuil de 5 %. Si  $p \leq 0,05$  : L'autocorrélation des résidus est significative, ce qui peut indiquer une violation de l'hypothèse d'indépendance des erreurs.

```
dwtest(model)
```

```
##
## Durbin-Watson test
##
## data: model
## DW = 2.0332, p-value = 0.1621
## alternative hypothesis: true autocorrelation is greater than 0
```

Dans notre cas, nous constatons clairement qu'il n'y a aucun problème, la statistique est très proche de 2.

## Limites du modèle

Notre modèle présente plusieurs limites qu'il est important de souligner. D'abord, il repose sur un échantillon de seulement 20 observations, ce qui limite la robustesse des résultats. De plus, nous avons supposé des relations linéaires entre les variables, ce qui peut ne pas refléter la réalité, car ces relations peuvent être plus complexes. Bien que nous ayons inclus plusieurs facteurs macroéconomiques, d'autres éléments importants, comme les politiques budgétaires ou des événements imprévus (crises économiques...), n'ont pas été pris en compte. De plus, les données utilisées étant nationales, elles ne capturent pas les différences régionales ou sectorielles. Même si le modèle respecte les principales hypothèses, des problèmes comme la corrélation entre certaines variables pourraient influencer les résultats. Cela montre qu'il serait utile d'explorer des modèles plus complexes ou d'utiliser des données plus précises pour améliorer l'analyse. Une autre limite importante réside dans le risque d'endogénéité. Certaines variables explicatives du modèle pourraient être corrélées avec l'erreur résiduelle, ce qui biaise les estimations des coefficients. Par exemple, des variables comme le PIB ou le chômage pourraient être influencées par des facteurs non observés (par exemple, des chocs structurels) qui affectent également la variable dépendante. Ce problème peut remettre en cause la validité de certaines inférences causales tirées du modèle.

## Conclusion

Cette étude a permis de mettre en évidence les principaux facteurs macroéconomiques influençant le déficit public en France au cours des 20 dernières années. Le taux de chômage, le taux d'épargne des ménages, le taux de change effectif, ainsi que les recettes fiscales nettes se révèlent être des variables significativement liées au déficit public. Une augmentation du chômage ou du taux d'épargne des ménages tend à aggraver le déficit, tandis qu'une appréciation de la monnaie ou une hausse des recettes fiscales semblent contribuer à sa réduction. Le PIB joue également un rôle positif, mais dans une moindre mesure, alors que l'inflation n'a pas montré de lien significatif avec le déficit.

Les implications politiques de cette étude sont donc claires, les décideurs publics peuvent agir sur certains facteurs économiques clés pour maîtriser le déficit public :

**Réduction du chômage** : Les résultats montrent qu'une réduction du taux de chômage serait une stratégie efficace pour améliorer la situation budgétaire. Les politiques d'emploi et les investissements favorisant la création d'emplois devraient donc être prioritaires.

**Réduire l'épargne excessive** : Le lien négatif entre l'épargne des ménages et le déficit public suggère que stimuler la consommation pourrait être une voie pour renforcer les recettes fiscales. Des incitations à la consommation, via des réductions d'impôts ou des politiques d'incitation à l'investissement, pourraient être envisagées.

**Renforcement des recettes fiscales** : L'effet des recettes fiscales sur le déficit souligne l'importance de politiques fiscales efficaces. Un effort pourrait être fait pour élargir l'assiette fiscale ou améliorer le recouvrement des impôts, sans pour autant nuire à la croissance.

**Gestion du taux de change** : L'effet positif du taux de change sur le déficit met en lumière l'importance des politiques monétaires. Bien que l'appréciation d'une monnaie puisse avoir des répercussions sur les exportations, elle semble réduire le déficit en augmentant le pouvoir d'achat national pour les biens importés.

**Croissance économique** : Bien que l'effet du PIB soit modéré dans cette étude, il reste significatif. Favoriser la croissance via des investissements publics dans les infrastructures, l'innovation, et l'éducation demeure une stratégie à long terme pour améliorer les finances publiques.

Le modèle explique bien les variations du déficit public, avec un  $R^2$  ajusté de 92,46 %, ce qui montre que les variables choisies sont pertinentes. Cependant, il est important de rester prudent dans l'interprétation des résultats car les facteurs économiques sont complexes et des événements exceptionnels, comme la crise de 2008 ou celle du COVID-19, ont eu un impact significatif sur les finances publiques, ce qui peut influencer la manière dont les résultats se traduisent dans la réalité.

En conclusion, cette analyse montre que certaines variables macroéconomiques ont un impact réel sur le déficit public. Bien que les résultats soient intéressants, il est évident que d'autres facteurs peuvent jouer un rôle important. Il serait donc utile d'aller plus loin dans l'analyse pour mieux comprendre ces dynamiques et peut-être proposer des solutions adaptées aux enjeux économiques futurs.