

# Économétrie des Données de Panel - Analyse Économétrique des Défaillances d'Entreprises en France

SCHNEIDER HUGO

7 Février 2025

## I. Introduction

Dans un contexte économique incertain, les faillites représentent souvent un indicateur avancé des cycles économiques. En effet, leur augmentation peut signaler une contraction économique imminente, ce qui rend leur analyse particulièrement pertinente dans la gestion des crises économiques. La politique économique actuelle en France, notamment les mesures de soutien aux entreprises pendant la crise sanitaire, peuvent ainsi jouer un rôle significatif en influençant les taux de faillite. Ainsi, il est nécessaire de comprendre les facteurs sous-jacents des faillites afin de permettre aux décideurs de mieux anticiper les besoins d'ajustements politiques.

L'objectif de cette étude est d'analyser les facteurs influençant les défaillances d'entreprises en France à l'aide de techniques économétriques avancées. Nous cherchons à déterminer l'impact du taux de chômage et du produit intérieur brut (PIB) sur le nombre de faillites d'entreprises sur une période allant de 1990 à 2024.

Nous utilisons plusieurs approches économétriques, notamment une régression OLS classique, un modèle de correction d'erreur (ECM) et un modèle ARDL (AutoRegressive Distributed Lag). Ces différentes approches permettent d'évaluer les relations de court et long terme entre les variables.

## II. Cadre Statistique

### 1. Données utilisées et source

Les données couvrent la période 1990 à 2024, avec une fréquence trimestrielle et proviennent de l'INSEE. Les variables analysées sont :

- **Défaillances d'entreprises** : Nombre de défaillances d'entreprises par date de jugement en France
- **Taux de chômage** : en pourcentage (chômage au sens du Bureau International du Travail (BIT) ; données corrigées des variations saisonnières, en moyenne trimestrielle.)
- **Produit Intérieur Brut** : en milliards d'euros

### 2. Statistiques descriptives

Analyse descriptive des variables :

Les statistiques descriptives montrent que le nombre de faillites d'entreprises en France varie considérablement (min : 371, max : 2399), avec une médiane de 1224 et une moyenne de 1290. Le taux de chômage oscille entre 7.1% et 10.7%, avec une moyenne de 9.0%. Le PIB, quant à lui, en croissance sur la période, évolue entre 389 et 650,9 milliards d'euros. Ces tendances suggèrent une corrélation potentielle entre la conjoncture économique et les faillites d'entreprises, ce qui motive l'analyse économétrique menée dans cette étude.

### 3. Tests de stationnarité\*\*

On applique ici le test de Dickey-Fuller augmenté (ADF) pour vérifier la stationnarité des variables. Si les variables ne sont pas stationnaires, on applique la différenciation.

Il est essentiel de s'assurer que les séries temporelles sont stationnaires avant d'effectuer une régression, car des séries non stationnaires peuvent conduire à des résultats biaisés et non fiables. La stationnarité garantit ainsi que les propriétés statistiques de la série, telles que la moyenne et la variance, ne changent pas au cours du temps. C'est pourquoi la différenciation des séries est appliquée ici pour rendre les données stationnaires et ainsi permettre une analyse économétrique robuste.

Les résultats des tests de stationnarité de Dickey-Fuller montrent que les séries `business_failures`, `Unemployment_rate` et `Gross_domestic_product` ne sont pas stationnaires au niveau (p-values > 0.05). Après différenciation, elles deviennent stationnaires avec des p-values inférieures à 0.01. Cela indique que ces variables sont intégrées d'ordre 1 (I(1)), ce qui justifie l'utilisation de modèles dynamiques comme l'ARDL ou l'ECM pour analyser leurs relations à long terme tout en capturant les dynamiques de court terme.

### 4. Résultats des Estimations et commentaires

```
##
## Comparaison des Modèles
## =====
##                                     Dependent variable:
##                                     -----
##                                     business_failures      d(business_failures)
##                                     OLS                  dynamic
##                                     (1)                  linear      (3)
## -----
## Unemployment_rate      83.155***
##                        (16.273)
##
## Gross_domestic_product -3.929***
##                        (0.220)
##
## L(business_failures, 1)                                -0.041*
##                                                         (0.021)
##
## d(Unemployment_rate)                                106.768***
##                                                         (31.897)
##                                                         93.034***
##                                                         (31.723)
##
## d(Gross_domestic_product)                            0.193
##                                                         (0.916)
##                                                         0.838
##                                                         (0.887)
##
## L(ecm, 1)                                              -0.131***
##                                                         (0.045)
##
## Constant      2,607.991***
##                (215.838)
##                -1.934
##                (8.127)
##                49.965*
##                (28.319)
## -----
## Test de Breusch-Pagan      0.00176018865676517      0.0140780882139195      0.574988712019971
## Statistique de Durbin-Watson      0.317385246069551      2.24358556332395      2.35601918424903
## Autocorrélation (1er lag)      0.804882249294482      -0.125583887972784      -0.190610846076701
```

## Observations	139	138	138
## R2	0.772	0.140	0.112
## Adjusted R2	0.769	0.121	0.093
## Residual Std. Error	191.845 (df = 136)	93.139 (df = 134)	94.647 (df = 134)
## F Statistic	230.347*** (df = 2; 136)	7.301*** (df = 3; 134)	5.659*** (df = 3; 134)
## =====			
## Note:			*p<0.1; **p<0.05; ***p<0.01

## Rapport d'analyse des modèles OLS, ARDL et ECM

Les modèles OLS, ARDL et ECM sont particulièrement adaptés à l'analyse des données de panel avec des tendances économiques. Le modèle OLS permet une estimation simple mais efficace des relations linéaires entre les variables, tandis que l'ARDL capture les relations à long terme et à court terme, en tenant compte de l'éventuelle présence de séries de données non stationnaires. Enfin, le modèle ECM (Error Correction Model) est utile pour ajuster rapidement les déséquilibres à court terme dans les séries économiques et refléter les ajustements à long terme.

Le tableau présente trois modèles d'analyse des faillites d'entreprises, avec une variable dépendante indiquant les `business_failures` (faillites d'entreprises) et la variation de cette variable dans différents cadres de régression. Nous allons examiner chaque modèle un par un et comparer les résultats pour en tirer les conclusions nécessaires.

L'objectif de cette analyse économétrique est ainsi d'évaluer les facteurs influençant les faillites d'entreprises en utilisant trois spécifications de modèles économétriques : une régression linéaire classique (OLS) et deux modèles dynamiques (régression dynamique et modèle dynamique avec correction d'erreur – ECM).

**Modèle OLS (Ordinary Least Squares) :** Régression linéaire simple des faillites d'entreprises en fonction des variables économiques clés (taux de chômage et PIB). **Modèle Dynamique :** Régression en prenant en compte les lags (décalages) des faillites d'entreprises, ce qui permet d'intégrer un effet de persistance dans les données. **Modèle Dynamique avec ECM (Error Correction Model) :** Ajout d'un terme de correction d'erreur pour modéliser l'ajustement à long terme des faillites d'entreprises par rapport aux variables économiques.

### 1. Modèle OLS (Régression Linéaire)

**Le modèle OLS**, basé sur la régression linéaire, a permis de démontrer une relation significative entre les faillites d'entreprises et les variables explicatives étudiées.

En effet, concernant le taux de chômage (`Unemployment_rate`), on observe un coefficient de 83.155, avec une erreur standard de 16.273, ce qui indique ici une relation positive et significative à 1% ( $p < 0.01$ ). Cela suggère qu'une augmentation du taux de chômage est fortement associée à une hausse des faillites d'entreprises, ce qui est en ligne avec les attentes théoriques. Dès lors, un taux de chômage plus élevé réduit la demande de produits et services, affectant ainsi la rentabilité des entreprises.

Concernant le Produit intérieur brut (`Gross_domestic_product`), le coefficient de -3.929, avec une erreur standard de 0.220, ce qui révèle une relation négative et significative à 1% ( $p < 0.01$ ). Ainsi, une augmentation du PIB est ainsi associée à une baisse des faillites, ce qui est intuitivement cohérent avec l'idée que la croissance économique soutient la viabilité des entreprises.

**Autocorrélation et Hétéroscédasticité :** Le test de Durbin-Watson indique une forte autocorrélation ( $DW = 0.317$ ), ce qui est préoccupant pour la validité du modèle. En outre, le test de Breusch-Pagan (0.0017) suggère la présence d'hétéroscédasticité, ce qui peut affecter l'efficacité des estimateurs. **R<sup>2</sup> et F-statistique :** Avec un  $R^2$  de 0.772, le modèle explique 77,2% de la variance des faillites d'entreprises. La statistique F (230.347) est élevée et significative, confirmant que le modèle global est pertinent.

## 2. Modèle Dynamique (ARDL)

**Le modèle dynamique** ajuste les faillites d'entreprises pour tenir compte de l'effet de persistance, c'est-à-dire de l'influence des faillites passées sur celles de la période en cours.

Concernant le Lag des faillites d'entreprises ( $L1(\text{business\_failures})$ ), le coefficient de -0.041, bien que significatif à 10% ( $p < 0.1$ ), cela suggère que les faillites passées ont un effet négatif modéré sur les faillites actuelles. Cela indique qu'une entreprise en difficulté a plus de chances de faire face à de nouvelles faillites dans le futur, mais l'effet est relativement faible.

Lorsque l'on s'intéresse à la variation du taux de chômage ( $d(\text{Unemployment\_rate})$ ), on observe un coefficient de 106.768, significatif à 1% ( $p < 0.01$ ), ce qui souligne une forte réaction des faillites d'entreprises aux changements dans le taux de chômage. Cette dynamique est plus marquée par rapport au modèle OLS.

Quant à la variation du PIB ( $d(\text{Gross\_domestic\_product})$ ), on observe un coefficient de 0.193, non significatif, ce qui indique que les variations du PIB n'ont pas d'impact significatif dans le modèle dynamique, ce qui pourrait être dû à la nature de la variable ou à l'inefficacité de la spécification dynamique dans ce cas.

**Autocorrélation et Hétéroscédasticité** : Le modèle dynamique présente un test de Breusch-Pagan de 0.014, ce qui suggère la présence d'hétéroscédasticité, mais dans une moindre mesure que le modèle OLS. Le test de Durbin-Watson (2.243) est proche de 2, indiquant l'absence d'autocorrélation des résidus, ce qui est un résultat favorable. **R<sup>2</sup> et F-statistique** : Avec un R<sup>2</sup> de 0.140, ce modèle explique seulement 14% de la variance des faillites, ce qui montre que sa capacité explicative est relativement faible par rapport au modèle OLS. Cependant, la statistique F (7.301) reste significative, indiquant la pertinence du modèle dans le cadre dynamique.

## 3. Modèle Dynamique avec ECM (Error Correction Model)

L'introduction d'un terme de correction d'erreur (ECM) dans ce modèle permet d'examiner l'ajustement des faillites d'entreprises à long terme par rapport aux variables économiques.

**ECM** ( $L(\text{ecm}, 1)$ ) : Le coefficient de -0.131 est significatif à 1% ( $p < 0.01$ ), ce qui indique qu'une part de l'écart par rapport à l'équilibre à long terme est corrigée au cours de la période suivante. Ce résultat suggère que, même après un choc, le système des faillites tend à revenir à un équilibre au fil du temps.

**Variation du taux de chômage et du PIB** : Comme dans le modèle dynamique, la variation du taux de chômage reste significative, avec un coefficient de 93.034, mais la variation du PIB n'a toujours pas d'effet significatif.

**Autocorrélation et Hétéroscédasticité** : Le test de Breusch-Pagan (0.575) montre qu'il n'y a pas de problème d'hétéroscédasticité. La statistique de Durbin-Watson (2.356) confirme qu'il n'y a pas d'autocorrélation des résidus, ce qui est un bon signe.

**R<sup>2</sup> et F-statistique** : Le R<sup>2</sup> de 0.112 indique une faible capacité explicative du modèle, ce qui en fait le modèle le moins performant parmi les trois. La statistique F (5.659) reste significative, mais la faiblesse de l'explication suggère une limitation dans la capture des dynamiques des faillites d'entreprises.

### Performance des Modèles :

**Modèle OLS** : Il reste le plus robuste et le plus efficace pour expliquer les faillites d'entreprises, avec une capacité explicative élevée (R<sup>2</sup> de 0.772). Toutefois, la présence d'autocorrélation et d'hétéroscédasticité doit être prise en compte. **Modèle ARDL** : Bien qu'il capture les effets de persistance des faillites, ce modèle explique seulement 14% de la variance, ce qui limite son utilité pratique pour prédire les faillites. L'effet du taux de chômage demeure prépondérant. **Modèle ECM** : Ce modèle permet d'inclure une dynamique d'ajustement à long terme, mais sa capacité explicative reste très faible (R<sup>2</sup> de 0.112). Cela montre que l'introduction de la correction d'erreur n'améliore pas substantiellement les prédictions des faillites d'entreprises.

**Recommandations** : En effet, il serait pertinent de combiner les éléments des trois modèles dans une spécification plus sophistiquée, en particulier en prenant en compte les problèmes d'autocorrélation et d'hétéroscédasticité dans les modèles OLS. De plus, la spécification dynamique pourrait être améliorée en tenant compte d'autres variables susceptibles d'influencer les faillites d'entreprises, telles que les taux d'intérêt ou l'investissement.

### III. Conclusion

Cette étude met en évidence les principaux déterminants des faillites d'entreprises en France. Nos résultats indiquent que : Une augmentation du taux de chômage entraîne une hausse significative des faillites d'entreprises. Le PIB a un effet négatif sur les faillites, ce qui est conforme à l'intuition économique : une croissance économique plus forte réduit le risque de faillite. Le modèle ECM suggère une relation de long terme entre ces variables, avec un ajustement partiel des déséquilibres au fil du temps. Les résultats des tests diagnostiques montrent la nécessité d'utiliser des erreurs robustes et des modèles dynamiques pour obtenir des estimations fiables.

Ces résultats suggèrent que les politiques économiques visant à stimuler la croissance du PIB et à réduire le chômage pourraient jouer un rôle clé dans la réduction du nombre de faillites d'entreprises en France. Une surveillance accrue des périodes de récession et des politiques de soutien aux entreprises en difficulté pourraient également atténuer les effets négatifs des cycles économiques.

En récapitulant les points essentiels, cette analyse met en lumière l'importance de comprendre les facteurs liés aux faillites pour mieux anticiper les fluctuations économiques. Une réflexion sur les implications politiques spécifiques pour le gouvernement français s'avère pertinente. Par exemple, les politiques de relance économique mises en place pendant la crise sanitaire ont pu contribuer à amortir les effets d'une récession, mais l'impact à long terme des faillites sur l'économie française nécessite une évaluation approfondie. Le rôle des politiques publiques, en particulier celles visant à soutenir l'activité économique, pourrait être déterminant dans la gestion de futures crises économiques.

### Références bibliographiques citées dans le document

- Greene, W. (2012). *Econometric Analysis*.
- Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data*.
- INSEE : Sources de données.

### Annexes

Code R :

```
# Chargement des bibliothèques
library(readxl)
library(dplyr)
library(tidyr)
library(lubridate)
library(zoo)
library(lmtest)
library(sandwich)
library(tseries)
library(ggplot2)
library(stargazer)
```

```

# Définition du répertoire de travail
setwd("/Users/schneiderhugo/Desktop")

# Chargement des données
data <- read_excel("Données.xlsx")

# Conversion des variables en numériques
data$business_failures <- as.numeric(data$business_failures)
data$Unemployment_rate <- as.numeric(data$Unemployment_rate)
data$Gross_domestic_product <- as.numeric(data$Gross_domestic_product)

# Vérification et correction de la colonne temporelle
data$Year_Trimestre <- as.yearqtr(as.character(data$Year_Trimestre), format = "%Y-T%q")

# Statistiques descriptives
summary(data)

```

```

##      Country      Year_Trimestre business_failures Unemployment_rate
## Length:139      Min.      :1990      Min.      : 371      Min.      : 7.100
## Class :character 1st Qu.:1999      1st Qu.:1058      1st Qu.: 8.050
## Mode  :character Median :2007      Median :1224      Median : 9.000
##                      Mean  :2007      Mean  :1290      Mean  : 9.009
##                      3rd Qu.:2016      3rd Qu.:1478      3rd Qu.:10.050
##                      Max.   :2024      Max.   :2399      Max.   :10.700
## Gross_domestic_product
## Min.      :389.0
## 1st Qu.:454.7
## Median :545.7
## Mean  :526.3
## 3rd Qu.:587.9
## Max.   :650.9

```

```

# Vérification de la stationnarité avec le test de Dickey-Fuller (ADF)
adf.test(data$business_failures)

```

```

##
## Augmented Dickey-Fuller Test
##
## data: data$business_failures
## Dickey-Fuller = -3.4021, Lag order = 5, p-value = 0.05714
## alternative hypothesis: stationary

```

```

adf.test(data$Unemployment_rate)

```

```

##
## Augmented Dickey-Fuller Test
##
## data: data$Unemployment_rate
## Dickey-Fuller = -2.2892, Lag order = 5, p-value = 0.4561
## alternative hypothesis: stationary

```

```
adf.test(data$Gross_domestic_product)
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: data$Gross_domestic_product  
## Dickey-Fuller = -2.3147, Lag order = 5, p-value = 0.4455  
## alternative hypothesis: stationary
```

```
# Si les séries ne sont pas stationnaires, on applique la différenciation  
data$business_failures_diff <- c(NA, diff(data$business_failures))  
data$Unemployment_rate_diff <- c(NA, diff(data$Unemployment_rate))  
data$Gross_domestic_product_diff <- c(NA, diff(data$Gross_domestic_product))  
  
# Vérification après différenciation  
adf.test(na.omit(data$business_failures_diff))
```

```
## Warning in adf.test(na.omit(data$business_failures_diff)): p-value smaller than  
## printed p-value
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: na.omit(data$business_failures_diff)  
## Dickey-Fuller = -4.682, Lag order = 5, p-value = 0.01  
## alternative hypothesis: stationary
```

```
adf.test(na.omit(data$Unemployment_rate_diff))
```

```
## Warning in adf.test(na.omit(data$Unemployment_rate_diff)): p-value smaller than  
## printed p-value
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: na.omit(data$Unemployment_rate_diff)  
## Dickey-Fuller = -4.8542, Lag order = 5, p-value = 0.01  
## alternative hypothesis: stationary
```

```
adf.test(na.omit(data$Gross_domestic_product_diff))
```

```
## Warning in adf.test(na.omit(data$Gross_domestic_product_diff)): p-value smaller  
## than printed p-value
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: na.omit(data$Gross_domestic_product_diff)  
## Dickey-Fuller = -6.2878, Lag order = 5, p-value = 0.01  
## alternative hypothesis: stationary
```

```
# Régression OLS classique
model_ols <- lm(business_failures ~ Unemployment_rate + Gross_domestic_product, data = data)
summary(model_ols)
```

```
##
## Call:
## lm(formula = business_failures ~ Unemployment_rate + Gross_domestic_product,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -739.86  -67.08    9.35   74.81  564.91
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2607.9912    215.8375   12.08 < 2e-16 ***
## Unemployment_rate     83.1548     16.2727    5.11 1.07e-06 ***
## Gross_domestic_product  -3.9289      0.2196  -17.89 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 191.8 on 136 degrees of freedom
## Multiple R-squared:  0.7721, Adjusted R-squared:  0.7687
## F-statistic: 230.3 on 2 and 136 DF, p-value: < 2.2e-16
```

```
# Correction des erreurs avec des erreurs robustes de Newey-West
coeftest(model_ols, vcov = NeweyWest(model_ols))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2607.9912    360.0465   7.2435 2.942e-11 ***
## Unemployment_rate     83.1548    105.7853   0.7861  0.43319
## Gross_domestic_product  -3.9289     1.7154  -2.2903  0.02354 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Vérification de l'autocorrélation des erreurs
dwtest(model_ols) # Durbin-Watson Test
```

```
##
## Durbin-Watson test
##
## data: model_ols
## DW = 0.31739, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
# Test d'hétéroscédasticité
bptest(model_ols) # Breusch-Pagan Test
```



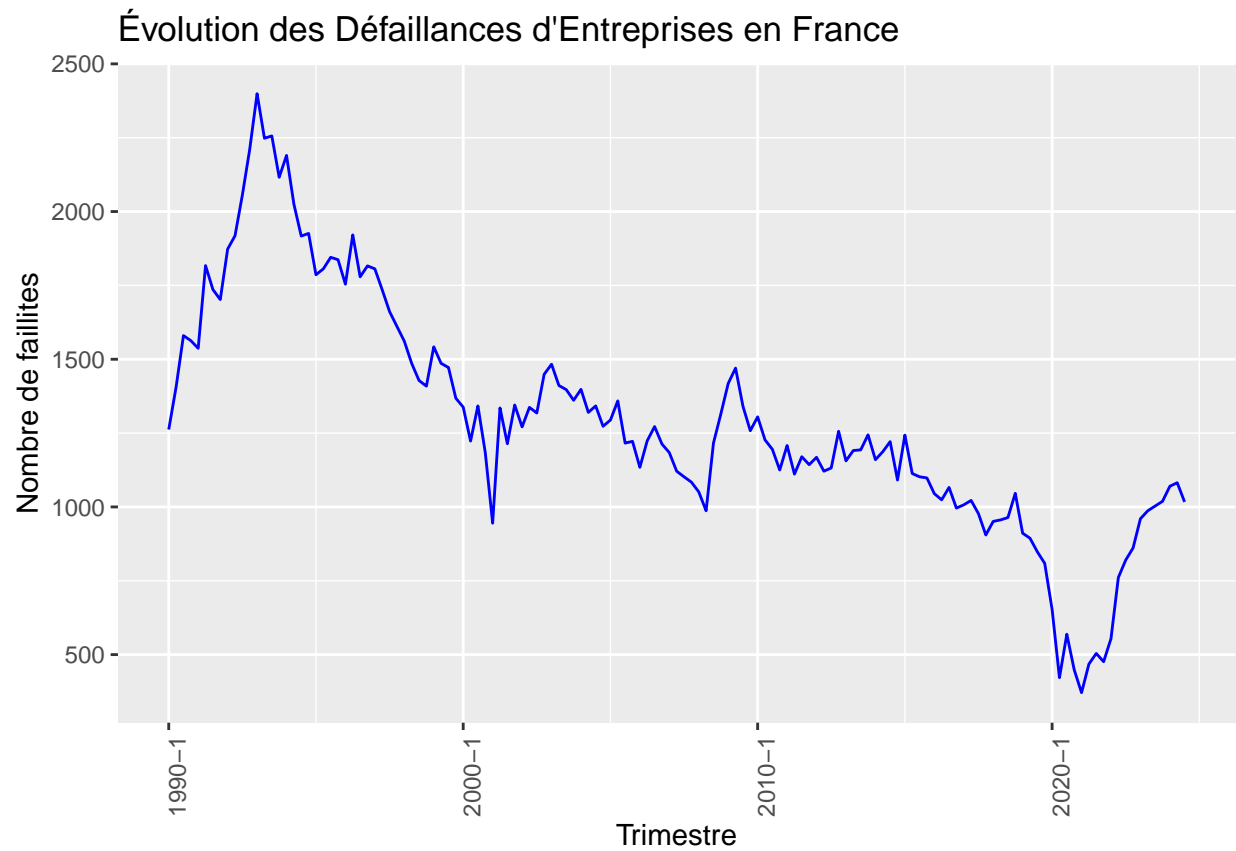
```
##
## studentized Breusch-Pagan test
##
## data: model_ols
## BP = 12.685, df = 2, p-value = 0.00176

# Si l'autocorrélation et l'hétéroscédasticité sont présentes, utiliser des erreurs robustes de White
coeftest(model_ols, vcov = vcovHC(model_ols, type = "HC1"))

##
## t test of coefficients:
##
##              Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)    2607.99121   228.52298   11.4124 < 2.2e-16 ***
## Unemployment_rate    83.15477    17.36937    4.7874 4.349e-06 ***
## Gross_domestic_product  -3.92888     0.25251  -15.5593 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

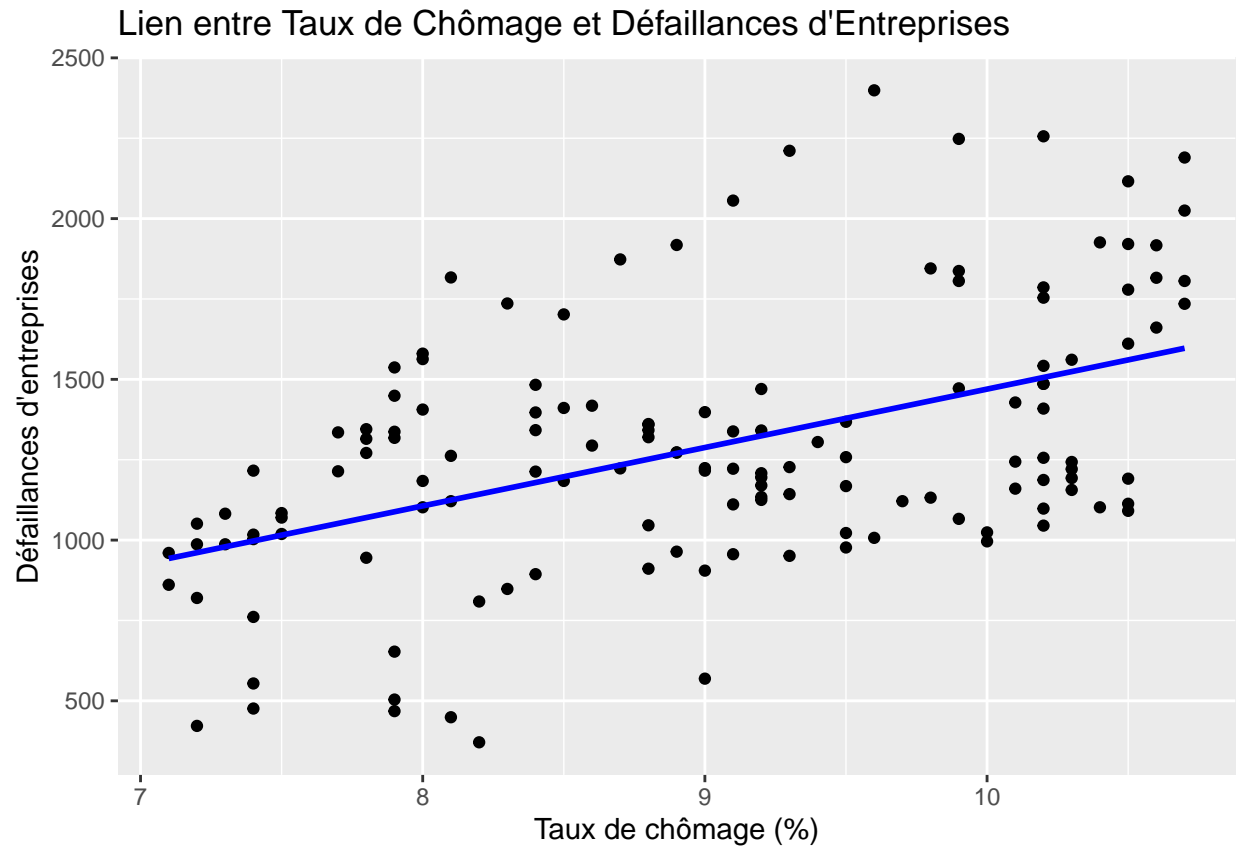
# Visualisation des variables
ggplot(data, aes(x = Year_Trimestre, y = business_failures, group = 1)) +
  geom_line(color = "blue") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  labs(title = "Évolution des Défaillances d'Entreprises en France",
       x = "Trimestre",
       y = "Nombre de faillites")

## Warning: The 'trans' argument of 'continuous_scale()' is deprecated as of ggplot2 3.5.0.
## i Please use the 'transform' argument instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```



```
ggplot(data, aes(x = Unemployment_rate, y = business_failures)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "blue") +  
  labs(title = "Lien entre Taux de Chômage et Défaillances d'Entreprises",  
        x = "Taux de chômage (%)",  
        y = "Défaillances d'entreprises")
```

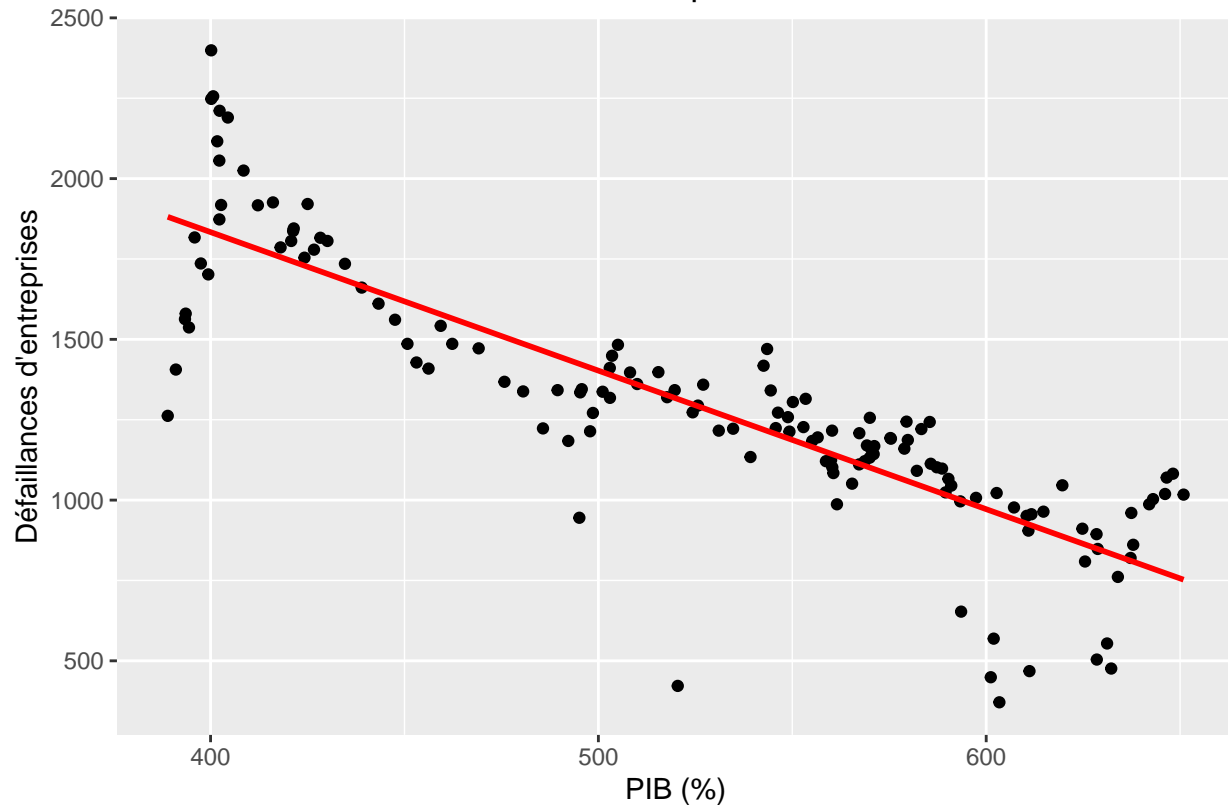
```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
ggplot(data, aes(x = Gross_domestic_product, y = business_failures)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Lien entre PIB et Défaillances d'Entreprises",  
        x = "PIB (%)",  
        y = "Défaillances d'entreprises")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

## Lien entre PIB et Défaillances d'Entreprises



```
str(data$Year_Trimestre) # Vérifie le type de la colonne
```

```
## 'yearqtr' num [1:139] 1990 Q1 1990 Q2 1990 Q3 1990 Q4 ...
```

```
table(data$Year_Trimestre) # Vérifie les valeurs uniques
```

```
##
## 1990 Q1 1990 Q2 1990 Q3 1990 Q4 1991 Q1 1991 Q2 1991 Q3 1991 Q4 1992 Q1 1992 Q2
##      1      1      1      1      1      1      1      1      1      1
## 1992 Q3 1992 Q4 1993 Q1 1993 Q2 1993 Q3 1993 Q4 1994 Q1 1994 Q2 1994 Q3 1994 Q4
##      1      1      1      1      1      1      1      1      1      1
## 1995 Q1 1995 Q2 1995 Q3 1995 Q4 1996 Q1 1996 Q2 1996 Q3 1996 Q4 1997 Q1 1997 Q2
##      1      1      1      1      1      1      1      1      1      1
## 1997 Q3 1997 Q4 1998 Q1 1998 Q2 1998 Q3 1998 Q4 1999 Q1 1999 Q2 1999 Q3 1999 Q4
##      1      1      1      1      1      1      1      1      1      1
## 2000 Q1 2000 Q2 2000 Q3 2000 Q4 2001 Q1 2001 Q2 2001 Q3 2001 Q4 2002 Q1 2002 Q2
##      1      1      1      1      1      1      1      1      1      1
## 2002 Q3 2002 Q4 2003 Q1 2003 Q2 2003 Q3 2003 Q4 2004 Q1 2004 Q2 2004 Q3 2004 Q4
##      1      1      1      1      1      1      1      1      1      1
## 2005 Q1 2005 Q2 2005 Q3 2005 Q4 2006 Q1 2006 Q2 2006 Q3 2006 Q4 2007 Q1 2007 Q2
##      1      1      1      1      1      1      1      1      1      1
## 2007 Q3 2007 Q4 2008 Q1 2008 Q2 2008 Q3 2008 Q4 2009 Q1 2009 Q2 2009 Q3 2009 Q4
##      1      1      1      1      1      1      1      1      1      1
## 2010 Q1 2010 Q2 2010 Q3 2010 Q4 2011 Q1 2011 Q2 2011 Q3 2011 Q4 2012 Q1 2012 Q2
```

```
##      1      1      1      1      1      1      1      1      1      1
## 2012 Q3 2012 Q4 2013 Q1 2013 Q2 2013 Q3 2013 Q4 2014 Q1 2014 Q2 2014 Q3 2014 Q4
##      1      1      1      1      1      1      1      1      1      1
## 2015 Q1 2015 Q2 2015 Q3 2015 Q4 2016 Q1 2016 Q2 2016 Q3 2016 Q4 2017 Q1 2017 Q2
##      1      1      1      1      1      1      1      1      1      1
## 2017 Q3 2017 Q4 2018 Q1 2018 Q2 2018 Q3 2018 Q4 2019 Q1 2019 Q2 2019 Q3 2019 Q4
##      1      1      1      1      1      1      1      1      1      1
## 2020 Q1 2020 Q2 2020 Q3 2020 Q4 2021 Q1 2021 Q2 2021 Q3 2021 Q4 2022 Q1 2022 Q2
##      1      1      1      1      1      1      1      1      1      1
## 2022 Q3 2022 Q4 2023 Q1 2023 Q2 2023 Q3 2023 Q4 2024 Q1 2024 Q2 2024 Q3
##      1      1      1      1      1      1      1      1      1
```

```
data_panel <- pdata.frame(data, index = c("Country", "Year_Trimestre"))
```

```
# Estimation de la relation de long terme
```

```
long_run_model <- lm(business_failures ~ Unemployment_rate + Gross_domestic_product, data = data)
```

```
data$ecm <- residuals(long_run_model) # Création du terme d'erreur
```

```
# Ajout d'ecm à 'data_zoo'
```

```
data_zoo$ecm <- coredata(data$ecm) # S'assurer que 'ecm' est bien aligné
```

```
# Modèle ECM avec 'dynlm()'
```

```
model_ecm <- dynlm(d(business_failures) ~ d(Unemployment_rate) + d(Gross_domestic_product) + L(ecm, 1),
```

```
# Résumé des résultats
```

```
summary(model_ecm)
```

```
##
```

```
## Time series regression with "zoo" data:
```

```
## Start = 1990 Q2, End = 2024 Q3
```

```
##
```

```
## Call:
```

```
## dynlm(formula = d(business_failures) ~ d(Unemployment_rate) +
```

```
##      d(Gross_domestic_product) + L(ecm, 1), data = data_zoo)
```

```
##
```

```
## Residuals:
```

```
##      Min      1Q  Median      3Q      Max
```

```
## -236.64 -68.73  -3.98   52.25  354.48
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)    -1.93422     8.12746  -0.238  0.81226
```

```
## d(Unemployment_rate)  106.76786    31.89723   3.347  0.00106 **
```

```
## d(Gross_domestic_product)  0.19318     0.91632   0.211  0.83335
```

```
## L(ecm, 1)        -0.13123     0.04533  -2.895  0.00443 **
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Residual standard error: 93.14 on 134 degrees of freedom
```

```
## Multiple R-squared:  0.1405, Adjusted R-squared:  0.1213
```

```
## F-statistic: 7.301 on 3 and 134 DF, p-value: 0.0001428
```

```
# Comparaison avec un modèle ARDL si nécessaire
```

```
# Définition des séries temporelles avec index trimestriel
```

```
data_zoo <- zoo(data[, c("business_failures", "Unemployment_rate", "Gross_domestic_product")],
               order.by = as.yearqtr(data$Year_Trimestre, format = "%Y-T%q"))
data_zoo
```

##	business_failures	Unemployment_rate	Gross_domestic_product
## 1990 Q1	1262	8.1	388.956
## 1990 Q2	1406	8.0	391.046
## 1990 Q3	1580	8.0	393.587
## 1990 Q4	1563	8.0	393.398
## 1991 Q1	1537	7.9	394.433
## 1991 Q2	1817	8.1	395.890
## 1991 Q3	1736	8.3	397.495
## 1991 Q4	1702	8.5	399.413
## 1992 Q1	1873	8.7	402.270
## 1992 Q2	1918	8.9	402.721
## 1992 Q3	2056	9.1	402.258
## 1992 Q4	2211	9.3	402.334
## 1993 Q1	2399	9.6	400.161
## 1993 Q2	2248	9.9	400.180
## 1993 Q3	2256	10.2	400.670
## 1993 Q4	2116	10.5	401.724
## 1994 Q1	2190	10.7	404.451
## 1994 Q2	2025	10.7	408.509
## 1994 Q3	1917	10.6	412.199
## 1994 Q4	1926	10.4	416.085
## 1995 Q1	1786	10.2	418.051
## 1995 Q2	1806	9.9	420.792
## 1995 Q3	1845	9.8	421.449
## 1995 Q4	1837	9.9	421.283
## 1996 Q1	1754	10.2	424.211
## 1996 Q2	1921	10.5	425.029
## 1996 Q3	1779	10.5	426.670
## 1996 Q4	1816	10.6	428.311
## 1997 Q1	1806	10.7	430.140
## 1997 Q2	1735	10.7	434.654
## 1997 Q3	1661	10.6	438.976
## 1997 Q4	1611	10.5	443.314
## 1998 Q1	1561	10.3	447.564
## 1998 Q2	1486	10.2	450.769
## 1998 Q3	1428	10.1	453.100
## 1998 Q4	1409	10.2	456.206
## 1999 Q1	1542	10.2	459.345
## 1999 Q2	1486	10.2	462.324
## 1999 Q3	1472	9.9	469.134
## 1999 Q4	1368	9.5	475.778
## 2000 Q1	1338	9.1	480.597
## 2000 Q2	1223	8.7	485.715
## 2000 Q3	1342	8.4	489.480
## 2000 Q4	1184	8.0	492.238
## 2001 Q1	945	7.8	495.091

## 2001 Q2	1335	7.7	495.324
## 2001 Q3	1214	7.7	497.864
## 2001 Q4	1345	7.8	495.719
## 2002 Q1	1271	7.8	498.593
## 2002 Q2	1337	7.9	501.104
## 2002 Q3	1318	7.9	502.989
## 2002 Q4	1449	7.9	503.533
## 2003 Q1	1483	8.4	505.044
## 2003 Q2	1411	8.5	502.946
## 2003 Q3	1397	8.4	508.166
## 2003 Q4	1361	8.8	510.004
## 2004 Q1	1398	9.0	515.474
## 2004 Q2	1320	8.8	517.723
## 2004 Q3	1342	8.8	519.681
## 2004 Q4	1273	8.9	524.317
## 2005 Q1	1294	8.6	525.676
## 2005 Q2	1359	8.8	527.027
## 2005 Q3	1216	9.0	531.049
## 2005 Q4	1222	9.1	534.742
## 2006 Q1	1134	9.2	539.223
## 2006 Q2	1224	9.0	545.740
## 2006 Q3	1272	8.9	546.310
## 2006 Q4	1213	8.4	549.267
## 2007 Q1	1184	8.5	555.142
## 2007 Q2	1121	8.1	558.680
## 2007 Q3	1102	8.0	560.325
## 2007 Q4	1084	7.5	560.597
## 2008 Q1	1051	7.2	565.427
## 2008 Q2	987	7.3	561.522
## 2008 Q3	1216	7.4	560.260
## 2008 Q4	1315	7.8	553.451
## 2009 Q1	1418	8.6	542.605
## 2009 Q2	1470	9.2	543.511
## 2009 Q3	1341	9.2	544.450
## 2009 Q4	1258	9.5	548.871
## 2010 Q1	1305	9.4	550.153
## 2010 Q2	1227	9.3	552.881
## 2010 Q3	1195	9.2	556.557
## 2010 Q4	1125	9.2	560.014
## 2011 Q1	1208	9.2	567.289
## 2011 Q2	1111	9.1	567.223
## 2011 Q3	1170	9.2	569.250
## 2011 Q4	1143	9.3	570.990
## 2012 Q1	1168	9.5	571.147
## 2012 Q2	1121	9.7	568.676
## 2012 Q3	1132	9.8	569.962
## 2012 Q4	1256	10.2	570.001
## 2013 Q1	1156	10.3	570.455
## 2013 Q2	1191	10.5	575.437
## 2013 Q3	1193	10.3	575.297
## 2013 Q4	1244	10.1	579.459
## 2014 Q1	1160	10.1	578.904
## 2014 Q2	1187	10.2	579.794
## 2014 Q3	1221	10.3	583.288

## 2014 Q4	1091	10.5	582.123
## 2015 Q1	1243	10.3	585.440
## 2015 Q2	1113	10.5	585.709
## 2015 Q3	1102	10.4	587.265
## 2015 Q4	1098	10.2	588.619
## 2016 Q1	1045	10.2	590.998
## 2016 Q2	1024	10.0	589.529
## 2016 Q3	1066	9.9	590.262
## 2016 Q4	996	10.0	593.330
## 2017 Q1	1007	9.6	597.380
## 2017 Q2	1022	9.5	602.703
## 2017 Q3	977	9.5	607.169
## 2017 Q4	905	9.0	610.898
## 2018 Q1	951	9.3	610.447
## 2018 Q2	956	9.1	611.669
## 2018 Q3	964	8.9	614.768
## 2018 Q4	1046	8.8	619.644
## 2019 Q1	911	8.8	624.805
## 2019 Q2	894	8.4	628.461
## 2019 Q3	848	8.3	628.763
## 2019 Q4	809	8.2	625.486
## 2020 Q1	653	7.9	593.537
## 2020 Q2	422	7.2	520.465
## 2020 Q3	569	9.0	601.940
## 2020 Q4	449	8.1	601.186
## 2021 Q1	371	8.2	603.401
## 2021 Q2	468	7.9	611.156
## 2021 Q3	504	7.9	628.525
## 2021 Q4	476	7.4	632.245
## 2022 Q1	554	7.4	631.179
## 2022 Q2	761	7.4	633.959
## 2022 Q3	820	7.2	637.234
## 2022 Q4	861	7.1	637.918
## 2023 Q1	960	7.1	637.429
## 2023 Q2	987	7.2	642.036
## 2023 Q3	1003	7.4	643.026
## 2023 Q4	1019	7.5	646.126
## 2024 Q1	1070	7.5	646.544
## 2024 Q2	1082	7.3	648.162
## 2024 Q3	1017	7.4	650.903

```
# Vérification de la conversion
head(data_zoo)
```

##	business_failures	Unemployment_rate	Gross_domestic_product
## 1990 Q1	1262	8.1	388.956
## 1990 Q2	1406	8.0	391.046
## 1990 Q3	1580	8.0	393.587
## 1990 Q4	1563	8.0	393.398
## 1991 Q1	1537	7.9	394.433
## 1991 Q2	1817	8.1	395.890



```
library(dynlm)
model_ardl <- dynlm(d(business_failures) ~ L(business_failures, 1) + d(Unemployment_rate) + d(Gross_domestic_product), data = data_zoo)
summary(model_ardl)
```

```
##
## Time series regression with "zoo" data:
## Start = 1990 Q2, End = 2024 Q3
##
## Call:
## dynlm(formula = d(business_failures) ~ L(business_failures, 1) +
##       d(Unemployment_rate) + d(Gross_domestic_product), data = data_zoo)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -224.29  -67.08   -4.00   48.98  387.82
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    49.96515    28.31891   1.764  0.07995 .
## L(business_failures, 1) -0.04093    0.02078  -1.970  0.05090 .
## d(Unemployment_rate)   93.03384   31.72338   2.933  0.00395 **
## d(Gross_domestic_product) 0.83795    0.88663   0.945  0.34631
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 94.65 on 134 degrees of freedom
## Multiple R-squared:  0.1124, Adjusted R-squared:  0.09258
## F-statistic: 5.659 on 3 and 134 DF,  p-value: 0.001111
```

```
# Résultats
stargazer(model_ols, model_ecm, model_ardl, type = "text", title = "Comparaison des Modèles")
```

```
##
## Comparaison des Modèles
## =====
##                               Dependent variable:
##                               -----
##                               business_failures      d(business_failures)
##                               OLS                    dynamic
##                               (1)                    linear      (3)
##                               -----
## Unemployment_rate           83.155***
##                               (16.273)
##
## Gross_domestic_product      -3.929***
##                               (0.220)
##
## L(business_failures, 1)                                -0.041*
##                                                         (0.021)
##
## d(Unemployment_rate)                                106.768***      93.034***
```

```

##                                     (31.897)                (31.723)
##
## d(Gross_domestic_product)         0.193                0.838
##                                     (0.916)                (0.887)
##
## L(ecm, 1)                          -0.131***
##                                     (0.045)
##
## Constant                2,607.991***          -1.934          49.965*
##                           (215.838)          (8.127)          (28.319)
##
## -----
## Observations                139                138                138
## R2                        0.772                0.140                0.112
## Adjusted R2              0.769                0.121                0.093
## Residual Std. Error      191.845 (df = 136)    93.139 (df = 134)    94.647 (df = 134)
## F Statistic              230.347*** (df = 2; 136) 7.301*** (df = 3; 134) 5.659*** (df = 3; 134)
## =====
## Note:                                     *p<0.1; **p<0.05; ***p<0.01

```