

Chapitre 3 - Modèle 2 : codes de réplcation et annexes

Table of contents

| | | |
|----------|---|-----------|
| 1 | Setup et chargement des packages | 1 |
| 2 | Construction des indicateurs de long-terme | 2 |
| 3 | Volatilité de la croissance | 5 |
| 3.1 | Régressions FE sur la volatilité (vol10) | 5 |
| 3.2 | Choix de l'indicateur d'informalité | 19 |
| 3.3 | Effets marginaux sur la volatilité | 24 |
| 4 | Persistance de la croissance | 45 |
| 5 | Visualisations | 49 |
| 6 | Robustesses et extensions | 79 |
| 6.1 | Panel GMM dynamique (Arellano–Bond one-step et two-step, puis System-GMM) | 79 |
| 6.2 | LSDV dynamique (PLM «Within» avec correction Arellano) | 84 |
| 6.3 | Anderson–Hsiao IV-LSDV | 86 |

1 Setup et chargement des packages

```
# Setup et chargement des packages
library(dplyr)
library(zoo)
library(fixest)
library(plm)
library(lmtest)
```

```
library(sandwich)
library(margins)
library(ggplot2)
library(mgcv)      # pour les modèles GAM
library(splines)   # pour les splines dans des lm/feols

# Réglages généraux de ggplot2 (thème clair)
theme_set(theme_minimal(base_size = 14))
```

Dans le script qui suit, le dataframe dont nous avons besoin au démarrage pour pouvoir relancer à l'identique tous les blocs est le suivant : “efpv_sustainability”. Téléchargeable depuis Github.

```
efpv_sustainability <- readRDS("/Users/.../Chap3/efpv_sustainability.rds")
df1_mod1 <- efpv_sustainability
```

2 Construction des indicateurs de long-terme

A savoir : rolling AR(1), half life, vol5 et vol10.

```
library(dplyr)
library(slides)
library(zoo)

df2 <- df1_mod1 %>%
  group_by(country) %>%
  arrange(year, .by_group = TRUE) %>%

  mutate(
    # 1) rolling AR(1) sur 10 ans
    ar1_10 = slide_dbl(
      DeltaY,
      ~ {
        tmp <- na.omit(.x)
        if (length(tmp) >= 5) {
          coef(lm(tmp[-1] ~ tmp[-length(tmp)]))[2]
        } else {
          NA_real_
        }
      },
```

```

    .before = 9,
    .complete = TRUE
  ),

  # 2) version "safe" : on ne conserve que les strictement entre 0 et 1
  phi_safe = if_else(
    ar1_10 > 0 & ar1_10 < 1,
    ar1_10,
    NA_real_
  ),

  # 3) demi-vie calculée une seule fois, sur phi_safe
  half_life10 = -log(2) / log(phi_safe)
) %>%

# 4) volatilités glissantes
mutate(
  vol5 = slide_dbl(DeltaY, ~ sd(.x, na.rm = TRUE), .before = 4, .complete
    ↪ = TRUE),
  vol10 = slide_dbl(DeltaY, ~ sd(.x, na.rm = TRUE), .before = 9, .complete
    ↪ = TRUE)
) %>%

ungroup() %>%
mutate(lag_DeltaY_1 = lag(DeltaY)) %>%
dplyr::select(
  country, year, income_group,
  DeltaY, lag_DeltaY_1, ar1_10, half_life10,
  vol5, vol10,
  GreenPotSharew, dge, self_emp_pct,
  democracy, deltaLnK, DeltaLnL,
  ays, gov_exp_pctgdp, open_pct,
  access_electricity, hhi_va
)

```

On a un warning : vient du fait que, pour certains pays et années, notre `ar1_10` est 0, et nous calculons quand même `log(ar1_10)` dans l'appel à `if_else()`. Or `log()` sur 0 ou un nombre négatif renvoie `NaN` et déclenche un warning. En plus, `dplyr::if_else()` est strict : il évalue les deux arguments (le « yes » et le « no ») sur l'intégralité du vecteur avant de renvoyer la valeur appropriée, donc même pour les cas où `!is.na(ar1_10) & ar1_10 > 0` est `FALSE`, R aura déjà tenté d'évaluer `log(ar1_10)` et émis ces `NaN`.

On propose un résumé descriptif de `df2` :

```

library(dplyr)
library(tidyr)

descr_df2 <- df2 %>%
  dplyr::select(
    DeltaY, ar1_10, half_life10,
    vol5, vol10,
    GreenPotSharew, dge, self_emp_pct,
    democracy, deltaLnK, DeltaLnL,
    ays, gov_exp_pctgdp, open_pct,
    access_electricity, hhi_va
  ) %>%
  dplyr::summarise(
    dplyr::across(
      dplyr::everything(),
      list(
        n      = ~ sum(!is.na(.)),
        mean   = ~ mean(., na.rm = TRUE),
        sd     = ~ sd(., na.rm = TRUE),
        min    = ~ min(., na.rm = TRUE),
        p10    = ~ quantile(., .10, na.rm = TRUE),
        p50    = ~ median(., na.rm = TRUE),
        p90    = ~ quantile(., .90, na.rm = TRUE),
        max    = ~ max(., na.rm = TRUE),
        na     = ~ sum(is.na(.))
      ),
      .names = "{.fn}_{.col}"
    )
  ) %>%
  tidyr::pivot_longer(
    cols      = dplyr::everything(),
    names_to  = c("stat", "variable"),
    names_sep = "_",
    values_to = "value"
  ) %>%
  tidyr::pivot_wider(
    names_from = stat,
    values_from = value
  ) %>%
  dplyr::select(variable, n, mean, sd, min, p10, p50, p90, max, na)

# Résultat
library(knitr)

```

```
kable(descr_df2, digits = 3, caption = "Résumé descriptif de df2")
```

3 Volatilité de la croissance

3.1 Régressions FE sur la volatilité (vol10)

1) On test pour vol10 (volatilité glissante 10 ans) :

```
library(plm)
library(sandwich)
library(lmtest)
library(car)          # pour vif()
library(effects)      # pour allEffects()
library(ggplot2)      # pour visualisations
library(sjPlot)       # pour export HTML
if (!require(visreg)) install.packages("visreg")
library(visreg)
library(stargazer)
library(dplyr)        # pour left_join

# 1) Création de l'objet panel
pdata <- pdata.frame(df2, index = c("country", "year"))

# 2) Estimation des modèles à effets fixes
mod1 <- plm(vol10 ~ DeltaY,
            data = pdata,
            model = "within",
            effect = "twoways")
mod2 <- plm(vol10 ~ DeltaY + dge + democracy,
            data = pdata,
            model = "within",
            effect = "twoways")
mod3 <- plm(vol10 ~ DeltaY + dge + democracy + GreenPotSharew + hhi_va,
            data = pdata,
            model = "within",
            effect = "twoways")

# 3) Tests de spécification
print(pFtest(mod2, mod3))
print(phptest(mod1, mod2))
```

```

print(bptest(mod3, studentize = TRUE))
print(pdwtest(mod3, alternative = "two.sided"))
print(vif(lm(vol10 ~ DeltaY + dge + democracy + GreenPotSharew + hhi_va,
             data = df2)))

# 4) Robustesse : double clustering (pays & temps)
vcov_dc <- vcovDC(mod3, type = "HC1")
se_dc    <- sqrt(diag(vcov_dc))

## 4.a. Table console/text avec p-values et astérisques
stargazer(mod3,
           se      = list(se_dc),
           type    = "text",
           title   = "mod3 - double clustering",
           covariate.labels =
             ↪ c("ΔY", "dge", "democracy", "GreenPotSharew", "hhi"),
           notes   = "Erreurs standards double-clustered (pays &
             ↪ temps).",
           report  = "vcsp",                      # v=valeurs,
             ↪ c="std.errors, s=stars, p=p-values",
           star.cutoffs = c(0.10, 0.05, 0.01),      # seuils pour *, **,
             ↪ ***
           digits   = 3)

# 5) Résultats
out_dir <- "/Users/.../Chap3/new_graph_chap3_mod2"

## 5.1. Effet partiel de dge (avec gg=TRUE)
p1 <- visreg(mod3, "dge",
             type = "conditional",
             gg   = TRUE,
             main = "Effet partiel de dge sur vol10",
             ylab = "vol10 (partiel)",
             xlab = "dge") +
  theme_minimal(base_size = 14) +
  theme(
    panel.background = element_rect(fill = "white"),
    plot.background  = element_rect(fill = "white")
  )

ggsave(filename = file.path(out_dir, "chap3mod2_effet_dge_vol10.png"),
        plot    = p1,

```

```

width = 6, height = 4, dpi = 300, bg = "white")

## 5.2. Effet partiel de democracy (avec gg=TRUE)
p2 <- visreg(mod3, "democracy",
             type = "conditional",
             gg = TRUE,
             main = "Effet partiel de democracy sur vol10",
             ylab = "vol10 (partiel)",
             xlab = "democracy") +
  theme_minimal(base_size = 14) +
  theme(
    panel.background = element_rect(fill = "white"),
    plot.background = element_rect(fill = "white")
  )

ggsave(filename = file.path(out_dir, "chap3mod2_effet_democracy_vol10.png"),
        plot = p2,
        width = 6, height = 4, dpi = 300, bg = "white")

## 5.3 Séries temporelles des résidus par pays
resid_pser <- residuals(mod3)
idx <- attr(resid_pser, "index")
resid_df <- data.frame(
  country = idx[[1]],
  year = as.numeric(as.character(idx[[2]])),
  resid_mod3 = as.numeric(resid_pser)
)
df2_with_resid <- df2 %>% left_join(resid_df, by=c("country","year"))

p3 <- ggplot(df2_with_resid, aes(x=year, y=resid_mod3, group=country)) +
  geom_line(alpha=0.2) +
  labs(title="Résidus de mod3 par pays", x="Année", y="Résidu") +
  theme_minimal(base_size=14) +
  theme(panel.background=element_rect(fill="white"),
        plot.background=element_rect(fill="white"))
ggsave(file.path(out_dir, "chap3mod2_residus_mod3.png"),
        plot=p3, width=6, height=4, dpi=300, bg="white")

# 6) Export des résultats (tables LaTeX & HTML)
## 6.a. Table LaTeX avec p-values et astérisques
stargazer(mod1, mod2, mod3,
           se = list(

```

```

sqrt(diag(vcovHC(mod1, type="HC1",
  ↪ cluster="group"))),
sqrt(diag(vcovHC(mod2, type="HC1",
  ↪ cluster="group"))),
sqrt(diag(vcovHC(mod3, type="HC1",
  ↪ cluster="group")))
),
type          = "latex",
file          = "table4_vol10.tex",
title         = "Tableau 4.1 - Volatilité (vol10)",
dep.var.labels = "vol10",
column.labels = c("(1) Base", "(2) + Macro", "(3) Complet"),
covariate.labels =
  ↪ c("ΔY", "dge", "democracy", "GreenPotSharew", "hhi"),
notes         = "Erreurs standards robustes (clustered par
  ↪ pays).",
report        = "vcsp",
star.cutoffs  = c(0.10, 0.05, 0.01),
digits        = 3)

## 6.b. Table HTML (tab_model)
tab_model(mod1, mod2, mod3,
  file      = "table4_vol10.html",
  dv.labels = "vol10",
  pred.labels = c("ΔY", "dge", "democracy", "GreenPotSharew", "hhi"),
  title     = "Tableau 4.1 - Volatilité (vol10)",
  show.p    = TRUE,
  p.style   = "stars")

```

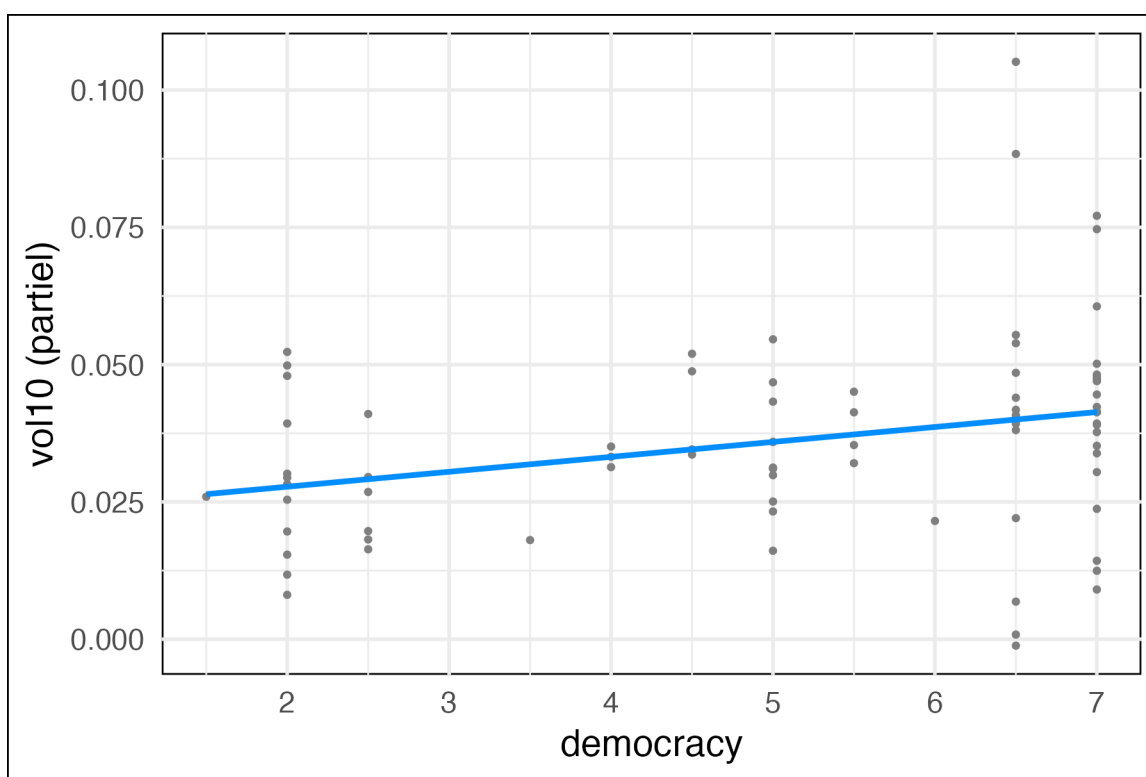
Tableau B.1 - Régression FE sur vol10

| Test / Variable | Statistique | Écart-type | p-value |
|---------------------------------|--|------------|-------------|
| F-twoways (mod2 vs mod3) | F = 0.4903 (df1 = -9, df2 = 55) | | NA |
| Hausman (FE vs RE) | $\chi^2 = 1.5507$ (df = 1) | | 0.213 |
| Breusch–Pagan | BP = 32.117 | | 5.63×10 *** |
| Durbin–Watson | DW = 1.3277 | | 0.00183*** |
| VIF | ΔY=1.08, dge=1.28, democracy=2.99, ... | | |

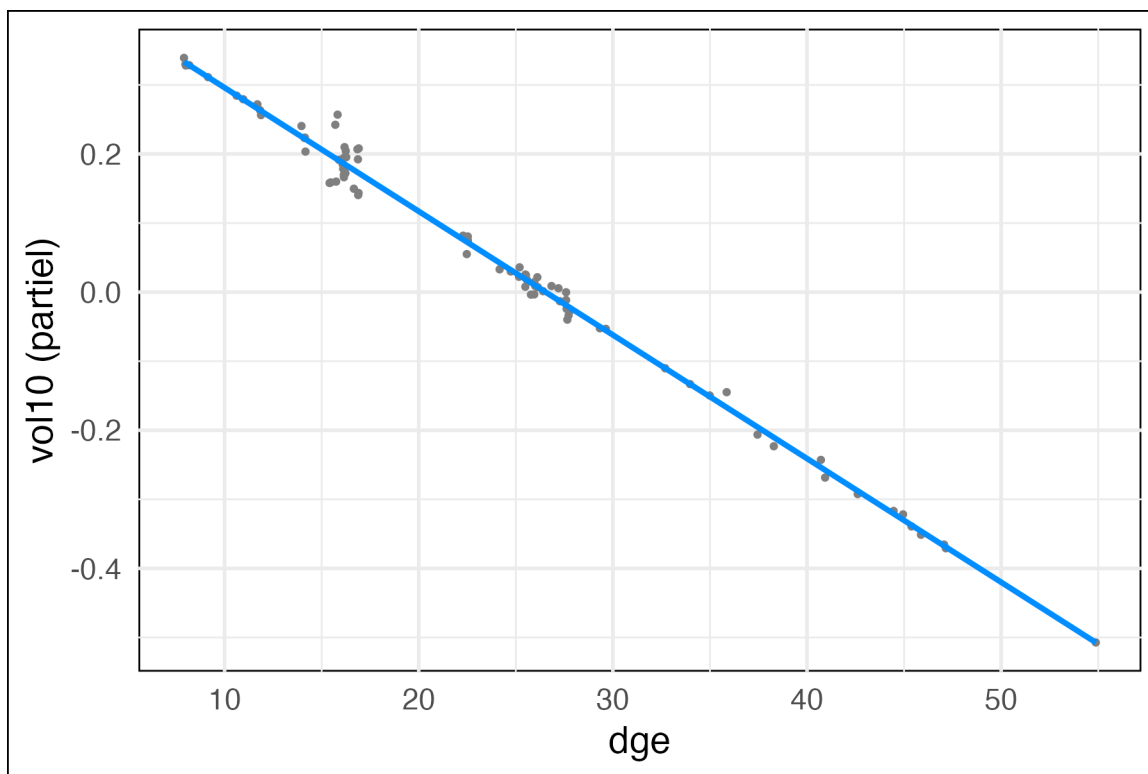
| Test / Variable | Statistique | Écart-type | p-value |
|-----------------------|-------------|------------|---------|
| — Régression | | | |
| mod3 (vol10) — | | | |
| ΔY | -0.105 | (0.069) | 0.127 |
| dge | -0.018 | (0.008) | 0.023** |
| democracy | 0.003 | (0.002) | 0.266 |
| GreenPotSharew | 0.100 | (0.262) | 0.704 |
| hhi | -0.666 | (0.359) | 0.064* |

Légende des astérisques : * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

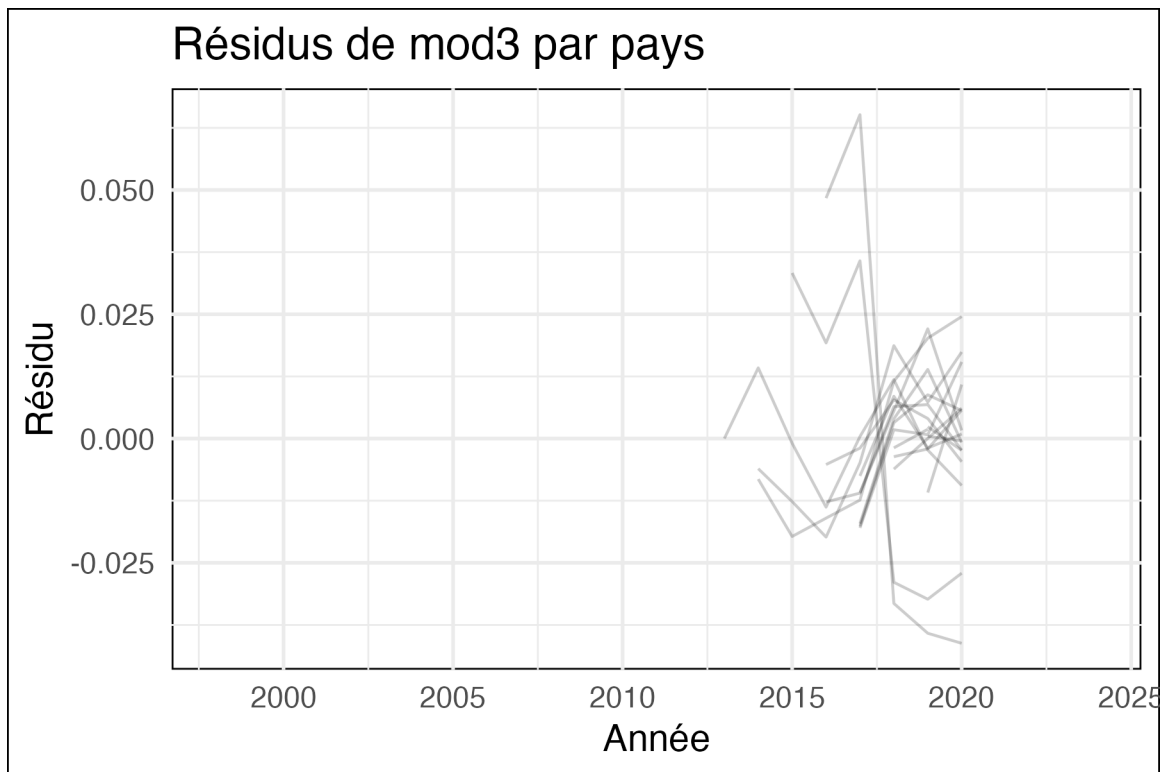
Graphique B.1 - vol10 (partiel) x democracy



Graphique B.2 - vol10 (partiel) x dge



Graphique B.3 - Résidus de mod3 par pays



2) On fait un test miroir pour vol5 cette fois (volatilité glissante sur 5 ans) :

```
library(plm)
library(sandwich)
library(lmtest)
library(car)          # pour vif()
library(effects)      # pour allEffects()
library(ggplot2)
library(sjPlot)       # pour export HTML
if (!require(visreg)) install.packages("visreg")
library(visreg)
library(stargazer)
library(dplyr)

# 1) Nommage des variables
names(df2)[names(df2) == "vol5"]           <- "vol"           # on utilise 'vol'
# ↪ pour plus de lisibilité
names(df2)[names(df2) == "ar1_10"]         <- "ar1"
names(df2)[names(df2) == "half_life10"]    <- "half"
names(df2)[names(df2) == "self_emp_pct"]   <- "self"
```

```

names(df2)[names(df2) == "gov_exp_pctgdp"] <- "gov"
names(df2)[names(df2) == "open_pct"] <- "open"
names(df2)[names(df2) == "access_electricity"] <- "access"
names(df2)[names(df2) == "hhi_va"] <- "hhi"

# 2) Création de l'objet panel
pdata5 <- pdata.frame(df2, index = c("country", "year"))

# 3) Estimation des modèles à effets fixes sur vol5
m1_5 <- plm(vol ~ DeltaY,
             data = pdata5,
             model = "within",
             effect = "twoways")

m2_5 <- plm(vol ~ DeltaY + dge + democracy,
             data = pdata5,
             model = "within",
             effect = "twoways")

m3_5 <- plm(vol ~ DeltaY + dge + democracy + GreenPotSharew + hhi,
             data = pdata5,
             model = "within",
             effect = "twoways")

# 4) Tests de spécification
pf_23_5 <- pFtest(m2_5, m3_5); print(pf_23_5)
haus_12_5 <- phptest(m1_5, m2_5); print(haus_12_5)
bp_m3_5 <- bptest(m3_5, studentize = TRUE); print(bp_m3_5)
dw_m3_5 <- pdwtest(m3_5, alternative = "two.sided"); print(dw_m3_5)
vif_m3_5 <- vif(lm(vol ~ DeltaY + dge + democracy + GreenPotSharew + hhi,
                  data = df2)); print(vif_m3_5)

# 5) Robustesse : double clustering (pays & temps)
vcov_dc5 <- vcovDC(m3_5, type = "HC1")
se_dc5 <- sqrt(diag(vcov_dc5))

stargazer(m3_5,
           se = list(se_dc5),
           type = "text",
           title = "Modèle 3 sur vol5 - double clustering",
           covariate.labels = c("ΔY", "dge", "democracy", "GreenPotSharew",
                                ↪ "hhi"),
           notes = "Erreurs standards double-clustered (pays &
                                ↪ temps).")

```

```

# 6) Résultats
out_dir5 <- "/Users/.../Chap3/new_graph_chap3_mod2"

## 6.1. Effet partiel de dge (gg=TRUE pour récupérer objet ggplot)
p1_5 <- visreg(m3_5, "dge",
               type = "conditional",
               gg    = TRUE,
               main  = "Effet partiel de dge sur vol5",
               ylab  = "vol5 (partiel)",
               xlab  = "dge") +
  theme_minimal(base_size = 14) +
  theme(panel.background=element_rect(fill="white"),
        plot.background=element_rect(fill="white"))

ggsave(filename = file.path(out_dir5, "chap3mod2_effet_dge_vol5.png"),
        plot    = p1_5, width = 6, height = 4, dpi = 300, bg = "white")

## 6.2. Effet partiel de democracy
p2_5 <- visreg(m3_5, "democracy",
               type = "conditional",
               gg    = TRUE,
               main  = "Effet partiel de democracy sur vol5",
               ylab  = "vol5 (partiel)",
               xlab  = "democracy") +
  theme_minimal(base_size = 14) +
  theme(panel.background=element_rect(fill="white"),
        plot.background=element_rect(fill="white"))

ggsave(filename = file.path(out_dir5, "chap3mod2_effet_democracy_vol5.png"),
        plot    = p2_5, width = 6, height = 4, dpi = 300, bg = "white")

## 6.3. Séries temporelles des résidus par pays
resid_pser5 <- residuals(m3_5)
idx5        <- attr(resid_pser5, "index")
resid_df5   <- data.frame(
  country    = idx5[[1]],
  year       = as.numeric(as.character(idx5[[2]])),
  resid_m3_5 = as.numeric(resid_pser5)
)
df2_with_resid5 <- df2 %>% left_join(resid_df5, by = c("country", "year"))

p3_5 <- ggplot(df2_with_resid5, aes(x = year, y = resid_m3_5, group =
  ↪ country)) +

```

```

geom_line(alpha = 0.2) +
labs(title="Résidus de m3_5 par pays", x="Année", y="Résidu") +
theme_minimal(base_size=14) +
theme(panel.background=element_rect(fill="white"),
      plot.background=element_rect(fill="white"))

ggsave(filename = file.path(out_dir5, "chap3mod2_residus_m3_vol5.png"),
      plot = p3_5, width = 6, height = 4, dpi = 300, bg = "white")

# 7) Export des résultats (tables LaTeX & HTML)
stargazer(m1_5, m2_5, m3_5,
          se = list(
            sqrt(diag(vcovHC(m1_5, type="HC1",
                              ↪ cluster="group"))),
            sqrt(diag(vcovHC(m2_5, type="HC1",
                              ↪ cluster="group"))),
            sqrt(diag(vcovHC(m3_5, type="HC1",
                              ↪ cluster="group"))),
          ),
          type = "latex",
          file = "table_vol5.tex",
          title = "Tableau - Volatilité vol5",
          dep.var.labels = "vol5",
          column.labels = c("(1) Base", "(2) + Macro", "(3) Complet"),
          covariate.labels =
            ↪ c("ΔY", "dge", "democracy", "GreenPotSharew", "hhi"),
          notes = "Erreurs standards clusterisées par pays.")

tab_model(m1_5, m2_5, m3_5,
          file = "table_vol5.html",
          dv.labels = "vol5",
          pred.labels = c("ΔY", "dge", "democracy", "GreenPotSharew", "hhi"),
          title = "Tableau - Volatilité vol5")

```

Tableau B.2 - Synthèse de tests de spécification

| Test / Estimation | Valeur | Détails / Remarques | |
|--|-------------------|---|---|
| F-test pour effets fixes (2-ways) | F = 1.8134 | df1 = -14 ; df2 = 209 ; <i>p-value</i> = NA | → Effets non significatifs (incomplet) F-test |
| Hausman Test (RE vs FE) | $\chi^2 = 1.2496$ | <i>p</i> = 0.2636 df = 1 ; H : modèle RE préférable ; H : FE inconsistant | → pas de rejet de H Breusch-Pagan (hétéroscédasticité) |
| | BP = 28.129 | <i>p</i> = 3.43e-05 *** df = 5 ; hétéroscédasticité | |

significative | | **Durbin-Watson (auto-corrélation)** | DW = 1.1636 ; $p = 1.22\text{e-}11$ *** |
Autocorrélation significative des erreurs idiosyncratiques |

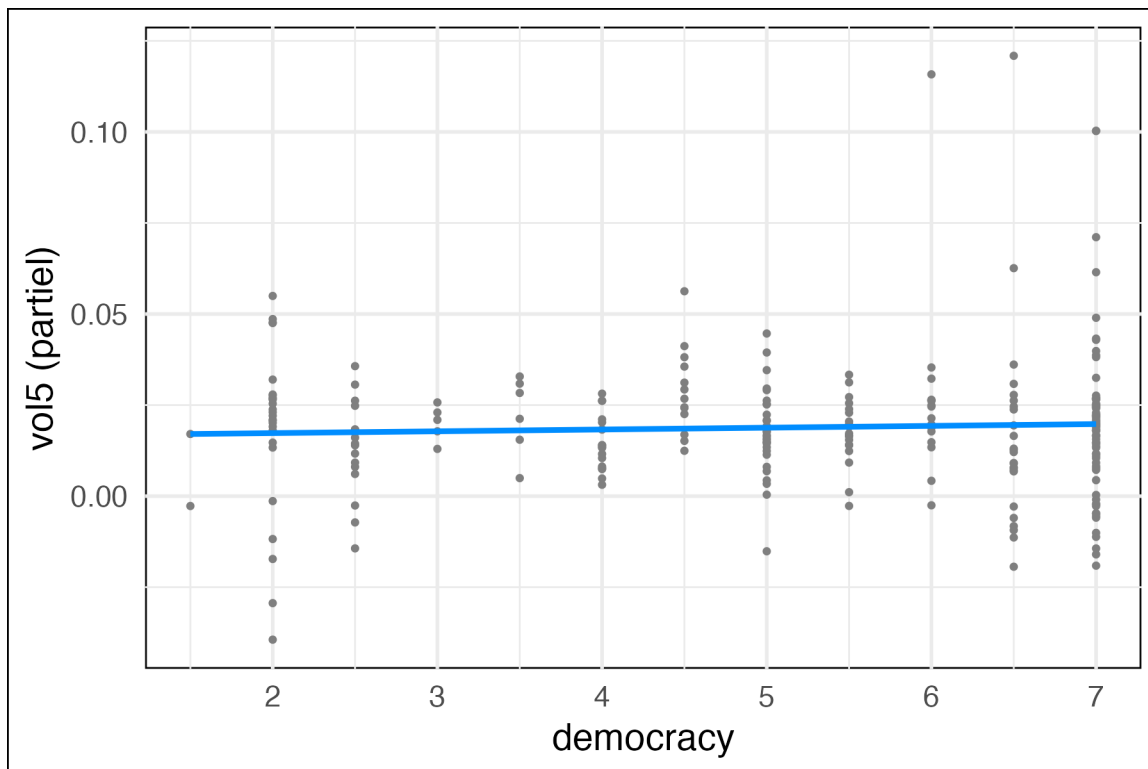
Tableau B.3 - Régression sur vol avec double clustering (pays)

| Variable indépendante | Coefficient | Erreur standard | Significativité | |
|-----------------------|-------------|-----------------|-----------------|----------------|
| ΔY | 0.039 | (0.154) | | dge |
| $p < 0.05$ | democracy | 0.0005 | (0.006) | GreenPotSharew |
| | | | | hhi |
| | | | | 0.870 |
| | | | | (0.565) |

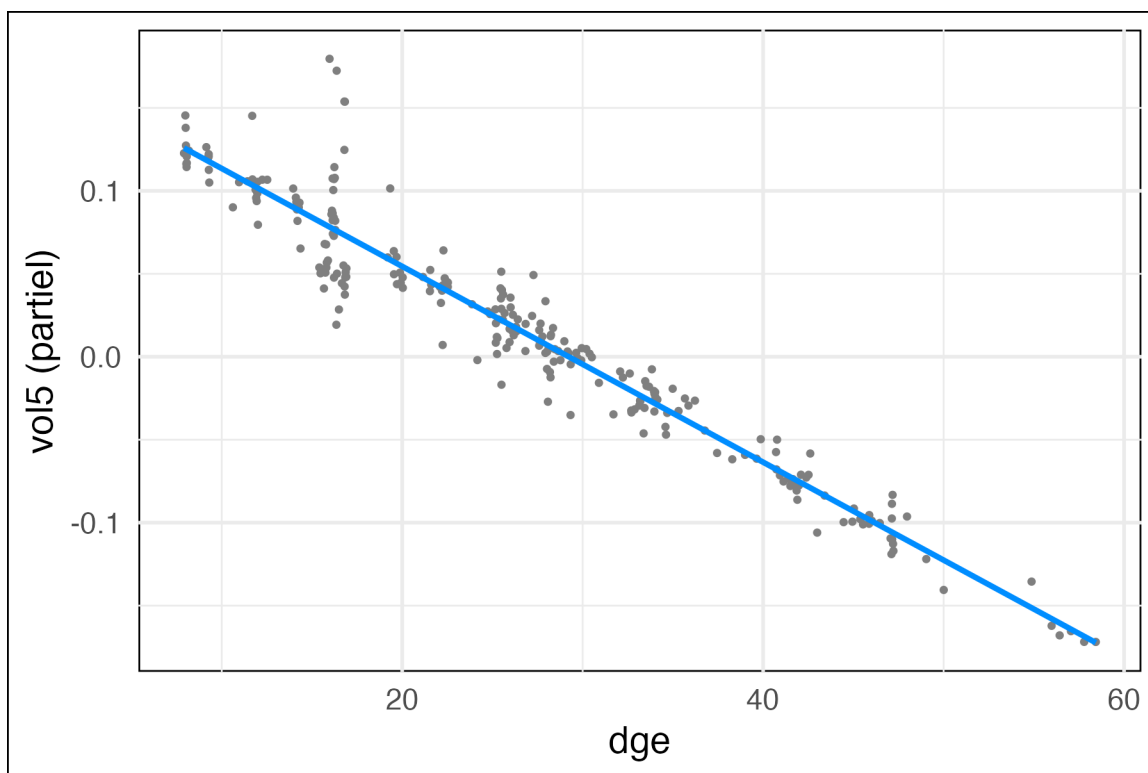
| Statistiques du modèle | Valeur |
|-------------------------|--------------------------|
| Observations | 258 |
| R ² | 0.156 |
| R ² ajusté | -0.112 |
| F-statistique | 7.205 *** (df = 5 ; 195) |
| Type d'erreurs standard | Double-clustered (pays) |

Colinéarité (VIF indicatif) | ΔY | dge | democracy | GreenPotSharew | hhi | | — | — | —
| ——— | — | | 1.14 | 1.23 | 2.48 | 2.01 | 2.41 |

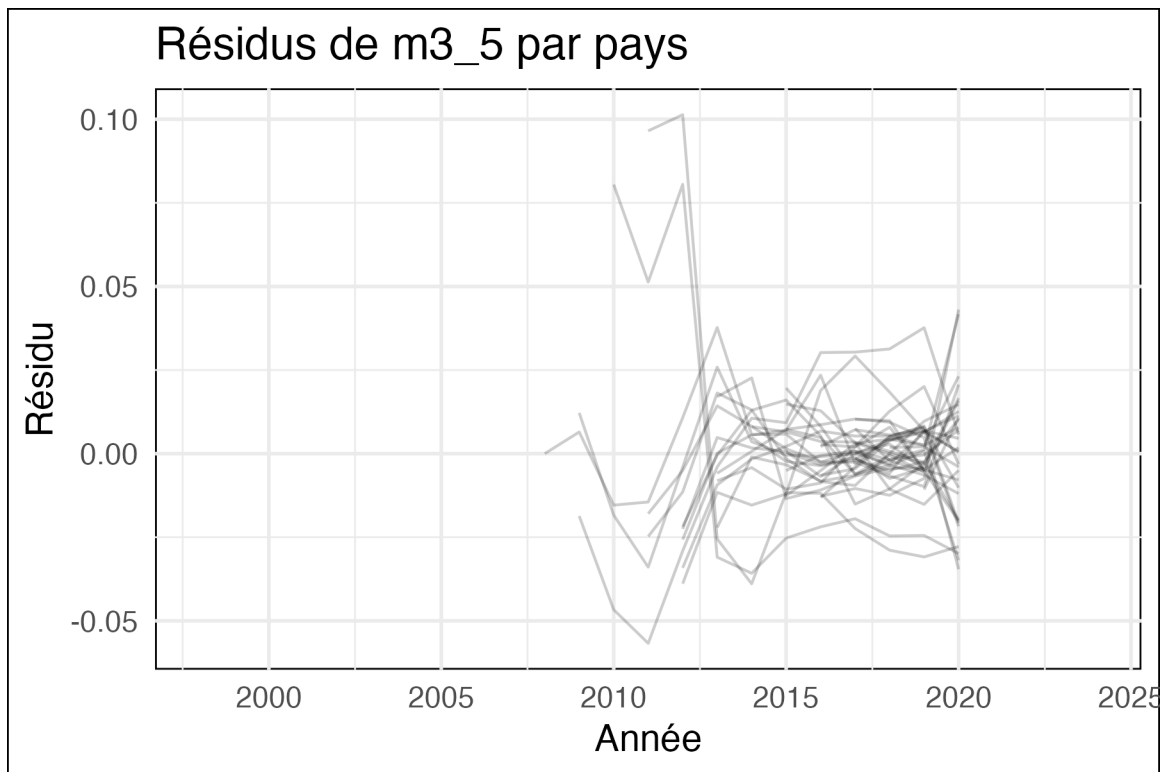
Graphique B.4 - vol5 (partiel) x democracy



Graphique B.5 - vol5 (partiel) x dge



Graphique B.6 - Résidus de m3_5 par pays



1) Erreur-type double-cluster (pays \times année) :

```
library(plm)
library(sandwich)
library(lmtest)

# 1) On estime explicitement mod3_10 et mod3_5

mod3_10 <- plm(vol10 ~ DeltaY + dge + democracy + GreenPotSharew + hhi_va,
  data = pdata, # pdata <- pdata.frame(df2,
    ↪ c("country", "year"))
  model = "within",
  effect = "twoways")

mod3_5 <- plm(vol5 ~ DeltaY + dge + democracy + GreenPotSharew + hhi_va,
  data = pdata,
  model = "within",
  effect = "twoways")

# 2) On double-clustering pour vol10
```

```
vcov_dc10 <- vcovDC(mod3_10, type = "HC1")
se_dc10 <- sqrt(diag(vcov_dc10))

cat("\n--- mod3_10 (vol10) double-clustered ---\n")
stargazer::stargazer(
  mod3_10,
  se = list(se_dc10),
  type = "text",
  title = "mod3_10 - double clustering", # ASCII only
  report = "vcsp",
  star.cutoffs = c(0.10, 0.05, 0.01),
  digits = 3
)

# 3) Double clustering pour vol5
vcov_dc5 <- vcovDC(mod3_5, type = "HC1")
se_dc5 <- sqrt(diag(vcov_dc5))

cat("\n--- mod3_5 (vol5) double-clustered ---\n")
stargazer::stargazer(
  mod3_5,
  se = list(se_dc5),
  type = "text",
  title = "mod3_5 - double clustering",
  report = "vcsp", # valeurs, std.errors, stars, p-values
  star.cutoffs = c(0.10, 0.05, 0.01),
  digits = 3
)
```

Tableau B.4 - comparatif des modèles : double clustering (pays&temps)

| Variables | vol10 (mod3_10) | vol5 (mod3_5) |
|----------------|---------------------------------------|---------------------------------------|
| ΔY | -0.105 (0.069) <i>p</i> = 0.127 | 0.039 (0.154) <i>p</i> = 0.799 |
| dge | -0.018 (0.008) ** <i>p</i> = 0.023 | -0.006 (0.003) ** <i>p</i> = 0.030 |
| democracy | 0.003 (0.002) <i>p</i> = 0.266 | 0.0005 (0.006) <i>p</i> = 0.937 |
| GreenPotSharew | 0.100 (0.262) <i>p</i> = 0.704 | 0.076 (0.094) <i>p</i> = 0.422 |

| Variables | vol10 (mod3_10) | vol5 (mod3_5) |
|-----------|------------------|---------------|
| hhi_va | -0.666 (0.359) * | 0.870 (0.565) |
| | $p = 0.064$ | $p = 0.124$ |

| Statistiques du modèle | vol10 | vol5 |
|------------------------|--------------------|-------------------------|
| Observations | 82 | 258 |
| R ² | 0.147 | 0.156 |
| R ² ajusté | -0.502 | -0.112 |
| F-statistique | 1.586 (df = 5; 46) | 7.205 *** (df = 5; 195) |

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Erreurs standards double-clustered (pays & années).

3.2 Choix de l'indicateur d'informalité

L'idée est de comparer plusieurs proxy de l'économie informelle dans nos régressions FE et de retenir celui qui est le plus significatif, améliore le plus les statistiques d'ajustement (R², AIC/BIC), et reste robuste au clustering double. Nous comparons séparément self_emp_pct, dge et mimic.

```
library(dplyr)
library(fixest)

# 1) Modèles pour vol10
## 1.1. Auto-emploi
mod10_self <- feols(
  vol10 ~ DeltaY + self + democracy + GreenPotSharew + hhi
  | country + year,
  data = df2,
  cluster = "country"
)

## 1.2. DGE
mod10_dge <- feols(
  vol10 ~ DeltaY + dge + democracy + GreenPotSharew + hhi
  | country + year,
  data = df2,
  cluster = "country"
)
```

```

# 2) Modèles pour vol5
## 2.1. Auto-emploi
mod5_self <- feols(
  vol ~ DeltaY + self + democracy + GreenPotSharew + hhi
  | country + year,
  data = df2,
  cluster = "country"
)

## 2.2. DGE
mod5_dge <- feols(
  vol ~ DeltaY + dge + democracy + GreenPotSharew + hhi
  | country + year,
  data = df2,
  cluster = "country"
)

# 3) Synthèse des résultats
## 3.1. Tableau pour vol10
etable(
  list(
    "Self-Emp" = mod10_self,
    "DGE" = mod10_dge
  ),
  dict = c(
    DeltaY = "ΔY",
    self = "Auto-emploi",
    dge = "DGE",
    democracy = "Démocratie",
    GreenPotSharew = "Part énergies vertes",
    hhi = "HHI VA"
  ),
  depvar = TRUE, # affiche "vol10" en en-tête
  digits = 3,
  title = "Volatilité 10-ans : Self-Emp vs DGE",
  fitstat = ~ r2
)

## 3.2. Tableau pour vol5
etable(
  list(

```

```

    "Self-Emp" = mod5_self,
    "DGE"      = mod5_dge
  ),
  dict = c(
    DeltaY      = "ΔY",
    self        = "Auto-emploi",
    dge         = "DGE",
    democracy   = "Démocratie",
    GreenPotSharew = "Part énergies vertes",
    hhi         = "HHI VA"
  ),
  depvar = TRUE,          # affiche "vol5"
  digits = 3,
  title   = "Volatilité 5-ans : Self-Emp vs DGE",
  fitstat = ~ r2
)

```

Tableau B.5 - comparaison des modèles Self-Emp vs DGE (panel vol10 et vol5)

| Variables | vol10 – Self-Emp | vol10 – DGE | vol5 – Self-Emp | vol5 – DGE |
|---------------------------|---------------------|----------------|--------------------|----------------|
| ΔY | 0.019 (0.033) | -0.105 (0.091) | 0.009 (0.061) | 0.039 (0.117) |
| Auto-emploi | 0.005 (0.004) | — | 0.003 (0.002) | — |
| Démocratie | -0.018 (0.010) | 0.003 (0.005) | -0.007 (0.007) | 0.0005 (0.006) |
| Partd'EFPV | 0.067 (0.083) | 0.100 (0.239) | 0.034 (0.074) | 0.076 (0.097) |
| HHIVA | -0.038 (0.181) | -0.666 (0.582) | 0.375 (0.297) | 0.870 (0.575) |
| DGE | — | -0.018 (0.009) | — | -0.006 (0.004) |
| Statistiques du modèle | vol10 – Self-Emp | vol10 – DGE | vol5 – Self-Emp | vol5 – DGE |
| Effets fixes pays | Oui | Oui | Oui | Oui |
| Effets fixes années | Oui | Oui | Oui | Oui |
| Clusterisation | Pays | Pays | Pays | Pays |
| R ² | 0.612 | 0.684 | 0.705 | 0.640 |

** $p < 0.1$; $p < 0.05$: * ; $p < 0.01$: ** ; $p < 0.001$: *** . Erreurs standards clusterisées par pays uniquement.*

On va rajouter des tests, commençons par extraire les Within R² et les AIC de nos quatre modèles :

```

library(dplyr)
library(plm)
library(lmtest)
library(sandwich)
library(tibble)
library(knitr)

# 1) suppression des obs. sans country ou year
df2_clean <- df2 %>%
  filter(!is.na(country), !is.na(year))

# 2) Identifier l'échantillon commun (sans NA pour toutes les variables
      ↪ utilisées)
vars <- c("vol10", "vol", "DeltaY", "self", "dge",
          "democracy", "GreenPotSharew", "hhi", "country", "year")

df_cmnn <- df2_clean %>%
  dplyr::select(all_of(vars)) %>%
  na.omit()

# 3) Passage en panel
pdata_cmnn <- pdata.frame(df_cmnn, index = c("country", "year"))

# 4) Estimation des 4 modèles à effets-fixes two-ways
mod10_self2 <- plm(vol10 ~ DeltaY + self,
                   data   = pdata_cmnn,
                   model   = "within",
                   effect  = "twoways")

mod10_dge2  <- plm(vol10 ~ DeltaY + dge,
                   data   = pdata_cmnn,
                   model   = "within",
                   effect  = "twoways")

mod5_self2  <- plm(vol    ~ DeltaY + self,
                   data   = pdata_cmnn,
                   model   = "within",
                   effect  = "twoways")

mod5_dge2   <- plm(vol    ~ DeltaY + dge,
                   data   = pdata_cmnn,
                   model   = "within",
                   effect  = "twoways")

```

```

# 5) Fonction pour calculer le R2 within d'un objet plm
r2_within_plm <- function(model) {
  mf      <- model.frame(model)
  yname   <- as.character(formula(model)[[2]])
  y       <- mf[[yname]]
  u       <- residuals(model)
  id      <- index(model)[[1]]
  ybar    <- ave(y, id)
  1 - sum(u^2, na.rm=TRUE) / sum((y - ybar)^2, na.rm=TRUE)
}

# 6) Extraction des R2 within
r2_10_self <- r2_within_plm(mod10_self2)
r2_10_dge  <- r2_within_plm(mod10_dge2)
r2_5_self  <- r2_within_plm(mod5_self2)
r2_5_dge   <- r2_within_plm(mod5_dge2)

# 7) Extraction des AIC (colonne "AIC" seulement)
aic_tab <- AIC(mod10_self2, mod10_dge2, mod5_self2, mod5_dge2)
aic_vec <- aic_tab[, "AIC"]

# 8) Construction du tableau de synthèse
results <- tibble(
  Modèle      = c("mod10_self2", "mod10_dge2", "mod5_self2", "mod5_dge2"),
  `R2 within` = c(r2_10_self, r2_10_dge, r2_5_self, r2_5_dge),
  AIC         = aic_vec
)

# 9) Affichage en Markdown
knitr::kable(
  results,
  digits      = c(0, 3, 4),
  caption     = "Comparaison des modèles sur échantillon commun"
)

```

Tableau B.6 - Synthétique des R² within et AIC pour les quatre modèles

| Modèle | R ² within | AIC |
|------------|-----------------------|-----------|
| vol10~self | 0.414 | -441.3160 |
| vol10~dge | 0.316 | -428.6438 |
| vol5~self | 0.713 | -559.7154 |

| Modèle | R ² within | AIC |
|----------|-----------------------|-----------|
| vol5~dge | 0.729 | -564.2574 |

Les résultats confirment une performance nettement supérieure des modèles appliqués à la volatilité de court terme (vol5) par rapport à ceux estimés sur la volatilité de moyen terme (vol10). En effet, la part de variance expliquée (R^2 within) est sensiblement plus élevée dans les modèles vol5 (entre 0.71 et 0.73), contre seulement 0.32 à 0.41 pour vol10. De même, les critères d'information d'Akaike (AIC) sont systématiquement plus faibles — donc meilleurs — pour les modèles vol5 (allant de -560 à -564) que pour vol10 (-428 à -441), indiquant un meilleur ajustement global. Ces résultats suggèrent que les variables explicatives mobilisées, notamment la part d'auto-emploi (`self_emp_pct`) ou l'indice synthétique d'informalité (`dge`), captent mieux les dynamiques de volatilité à 5 ans qu'à 10 ans. Dans le détail, la comparaison des deux spécifications montre une hiérarchie inversée selon l'horizon temporel retenu. À l'échelle décennale, le modèle intégrant la part d'auto-emploi (`mod10_self2`) présente de meilleures performances ($R^2 = 0.414$; $AIC = -441$) que celui mobilisant l'indice d'informalité (`mod10_dge2`, $R^2 = 0.316$; $AIC = -428$). À l'inverse, pour la volatilité quinquennale, c'est la spécification avec `dge` qui obtient les meilleurs résultats ($R^2 = 0.729$; $AIC = -564$), légèrement devant celle avec `self_emp_pct` ($R^2 = 0.713$; $AIC = -559$). Ainsi, si l'objectif est de maximiser la qualité de l'ajustement pour les dynamiques de volatilité à court terme, le modèle `mod5_dge2` s'impose comme le plus performant. En revanche, sur un horizon de 10 ans, la variable `self_emp_pct` conserve un avantage explicatif supérieur.

3.3 Effets marginaux sur la volatilité

- 1) On test GreenPotSharew x DGE pour mesurer l'impact du degré d'informalité sur l'impact du verdissement (vol10) :

```
library(dplyr)
library(fixest)
library(ggplot2)

# 1) Estimation du modèle avec interaction
mod_int10 <- feols(
  vol10 ~ DeltaY + democracy + hhi + GreenPotSharew * dge
  | country + year,
  data = df2,
  cluster = "country"
)

# 2) Grille de prédiction
```



```

dge_p <- quantile(df2$dge, probs = c(0.1,0.5,0.9), na.rm = TRUE)
m_DeltaY <- mean(df2$DeltaY, na.rm = TRUE)
m_dem <- mean(df2$democracy, na.rm = TRUE)
m_hhi <- mean(df2$hhi, na.rm = TRUE)

newdata <- expand.grid(
  GreenPotSharew = seq(
    min(df2$GreenPotSharew, na.rm = TRUE),
    max(df2$GreenPotSharew, na.rm = TRUE),
    length.out = 100
  ),
  dge = as.numeric(dge_p),
  DeltaY = m_DeltaY,
  democracy = m_dem,
  hhi = m_hhi
)

# 3) Extraire b et Vc cluster-robuste
b <- coef(mod_int10)
Vc <- vcov(mod_int10, cluster = "country")

# 4) Construction de la matrice de design (sans fixed-effects)
X_full <- model.matrix(
  ~ DeltaY + democracy + hhi + GreenPotSharew * dge,
  data = newdata
)

# 5) On garde et réordonne les colonnes selon les noms de b
b_names <- names(b)
X <- X_full[, b_names, drop = FALSE]

# 6) Calcul des fits et des SE
fits <- as.vector(X %>% b)
ses <- sqrt( rowSums((X %>% Vc[b_names,b_names]) * X) )

# 7) Préparation du data.frame de tracé
pred_int <- newdata %>%
  mutate(
    predicted = fits,
    conf.low = fits - 1.96 * ses,
    conf.high = fits + 1.96 * ses,
    group = factor(dge,
      levels = dge_p,

```

```

        labels = paste0(c("10e", "50e", "90e"), " pctlile"))
    )

# 8) Résultats
ggplot(pred_int, aes(x = GreenPotSharew, y = predicted,
                     colour = group, fill = group)) +
  geom_line(size = 1) +
  geom_ribbon(aes(ymin = conf.low, ymax = conf.high),
             alpha = 0.2, colour = NA) +
  labs(
    title     = "Effet marginal des EFPV sur vol10",
    subtitle  = "À trois niveaux de DGE (10e, 50e et 90e percentile)",
    x         = "Part énergies vertes (GreenPotSharew)",
    y         = "Volatilité 10-ans prédite (vol10)",
    colour    = "DGE", fill = "DGE"
  ) +
  theme_minimal(base_size = 14)

# 9) Résultats
p_int <- ggplot(pred_int, aes(x = GreenPotSharew, y = predicted,
                              colour = group, fill = group)) +
  geom_line(size = 1) +
  geom_ribbon(aes(ymin = conf.low, ymax = conf.high),
             alpha = 0.2, colour = NA) +
  labs(
    title     = "Effet marginal des EFPV sur vol10",
    subtitle  = "À trois niveaux de DGE (10e, 50e et 90e percentile)",
    x         = "Part énergies vertes (GreenPotSharew)",
    y         = "Volatilité 10-ans prédite (vol10)",
    colour    = "DGE", fill = "DGE"
  ) +
  theme_minimal(base_size = 14)

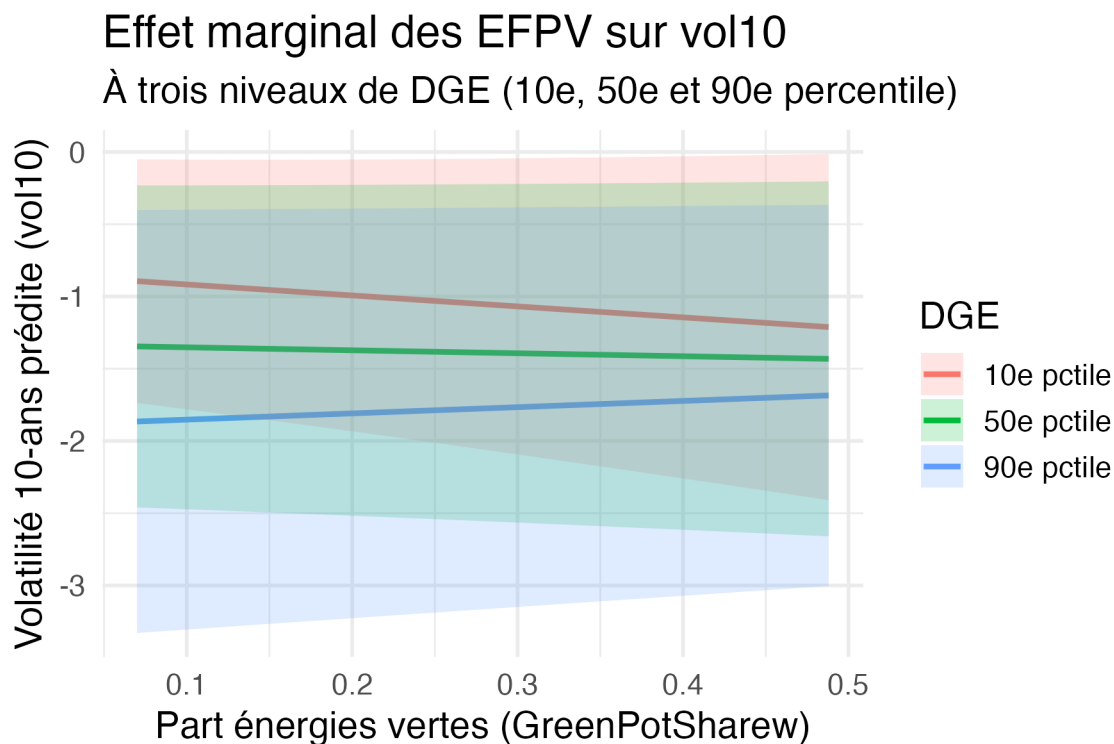
# 10) Téléchargement
out_dir_int <- "/Users/.../Chap3/int10"
if (!dir.exists(out_dir_int)) dir.create(out_dir_int, recursive = TRUE)

ggsave(
  filename = file.path(out_dir_int, "chap3mod2_effet_interaction_vol10.png"),
  plot     = p_int,
  width    = 6,
  height   = 4,
  dpi      = 300,

```

```
bg = "white"
)
```

Graphique B.7 - Effet marginal des EFPV sur vol10 à trois niveaux de percentile de DGE



L'analyse révèle que l'effet de la part d'emplois verts (**GreenPotSharew**) sur la volatilité de la croissance à moyen terme (**vol10**) dépend fortement du degré d'informalité du marché du travail (**DGE**). Trois profils types sont distingués, correspondant aux 10 , 50 et 90 percentiles de l'informalité. Dans les pays à **faible informalité** (10 percentile), l'effet du verdissement de l'emploi est globalement **stabilisateur** : plus la part d'emplois verts augmente, plus la volatilité à 10 ans tend à diminuer. Cet effet reste présent, bien que moins marqué, dans les contextes d'**informalité médiane** (50 percentile). En revanche, dans les pays à **forte informalité** (90 percentile), la relation s'inverse : une hausse de la part verte est associée à une **légère augmentation** de la volatilité macroéconomique, suggérant un effet potentiellement **déstabilisateur**. Les bandes de confiance associées à ces estimations sont relativement larges, notamment aux extrêmes du spectre de l'informalité, ce qui témoigne d'une incertitude plus grande lorsque les marchés du travail sont très peu ou très fortement formalisés. Ces résultats mettent en lumière un **effet conditionnel** de la transition verte : son potentiel stabilisateur

n'est pas universel. Il semble dépendre de la capacité des économies à absorber les transformations structurelles induites par le verdissement de l'emploi. Dans les contextes où l'informalité est élevée, les travailleurs peuvent rencontrer davantage de frictions à la reconversion vers des emplois verts protégés ou formels, ce qui augmente l'instabilité de la trajectoire économique. En termes de politiques publiques, cela suggère que le développement de l'emploi vert, s'il n'est pas accompagné de **mesures de formalisation, de formation professionnelle** ou de **protection sociale**, pourrait renforcer la vulnérabilité macroéconomique dans les pays du Sud ou à forte dualité productive. Soutenir la transition verte suppose donc une approche intégrée, sensible aux **conditions structurelles des marchés du travail**, pour éviter que cette transition ne devienne elle-même une source d'instabilité.

On test la robustesse de ces résultats, d'abord avec un test de l'hypothèse nulle "pente = 0" (Wald) :

```
library(fixest)

# 1) Modèle avec interaction
mod_int10 <- feols(
  vol10 ~ DeltaY + democracy + hhi + GreenPotSharew * dge
  | country + year,
  data = df2,
  cluster = "country"
)

# 2) test de Wald sur le terme d'interaction
wald(mod_int10, "GreenPotSharew:dge")

library(fixest)

# 3) Estimation du modèle avec interaction sur vol5 (ici nommé 'vol')
mod_int5 <- feols(
  vol ~ DeltaY + democracy + hhi + GreenPotSharew * dge
  | country + year,
  data = df2,
  cluster = "country"
)

# 3) Test de Wald pour H0 : coefficient de GreenPotSharew:dge = 0
wald5 <- wald(mod_int5, "GreenPotSharew:dge")
print(wald5)
```

L'analyse des termes d'interaction entre la part d'emplois verts et le niveau d'informalité du

marché du travail met en évidence un **effet conditionnel modéré à moyen terme**, mais **absent à court terme**. Sur l’horizon de 10 ans (modèle `vol10`), le test de Wald appliqué au seul coefficient d’interaction indique une statistique de 2.88 (ddl = 1, 45) pour une *p-value* 0.096. Ce résultat, **marginale­ment significatif au seuil de 10 %**, suggère que l’impact de la transition verte sur la volatilité dépend effectivement du degré d’informalité, bien que la robustesse statistique de cet effet reste limitée. En revanche, sur l’horizon de 5 ans (`vol5`), le même test donne une statistique de 0.65 avec une *p-value* de 0.42, indiquant **aucune signifi­cativité** : la part d’emplois verts ne semble pas modulée par l’informalité à court terme. Ces résultats conduisent à trois enseignements principaux. D’abord, l’**interaction GreenPotSharew × DGE** apparaît pertinente pour expliquer les dynamiques de **volatilité à moyen terme**, et mérite d’être conservée dans les modèles sur `vol10`. Ensuite, le **caractère temporel du mécanisme** se dessine clairement : les frictions liées à l’informalité — reconversions incomplètes, absence de filet social, rigidités du secteur informel — ne produisent leurs effets qu’à un horizon plus long. Enfin, à court terme, **aucune hétérogénéité significative** ne semble observable, ce qui justifie l’exclusion de l’interaction dans les spécifications sur `vol5`. Ces résultats ouvrent plusieurs pistes. Sur le plan empirique, ils incitent à renforcer la robustesse de l’interaction à moyen terme : tests de non-linéarité (splines, buckets), instrumentation alternative, ou calculs d’effets marginaux conditionnels (via `marginalEffects` ou `ggeffects`). Sur le plan interprétatif, ils appellent à **contextualiser la transition verte dans les structures productives existantes** : en contexte d’informalité élevée, elle pourrait provoquer à moyen terme une instabilité accrue, contrairement aux effets stabilisateurs attendus dans les économies plus formalisées.

Leave-one-country-out :

```
library(dplyr)
library(fixest)
library(tibble)

countries <- unique(df2$country)

res_loo <- tibble(
  country_exclu = character(),
  coef_int      = double(),
  se_int        = double()
)

for(c in countries) {
  df_sub <- df2 %>% filter(country != c)

  # 1) Vérifier qu'on a au moins 2 années différentes et une variance de
  ↪ vol10
  yrs <- n_distinct(df_sub$year)
```

```

varY <- var(df_sub$vol10, na.rm = TRUE)
if (yrs < 2 || is.na(varY) || varY == 0) {
  message(" Pays '", c, "' sauté (pas assez d'années ou vol10 constant)")
  next
}

# 2) On essaye d'estimer
m <- tryCatch({
  feols(
    vol10 ~ DeltaY + democracy + hhi + GreenPotSharew * dge
    | country + year,
    data = df_sub,
    cluster = "country"
  )
}, error = function(e) {
  message(" Estimation pour '", c, "' a échoué : ", e$message)
  return(NULL)
})
if (is.null(m)) next

# 3) On récupère le coefficient et SE de l'interaction
ct <- coef(m)[["GreenPotSharew:dge"]]
vt <- vcov(m, cluster = "country")[ "GreenPotSharew:dge",
↪ "GreenPotSharew:dge" ]
se <- sqrt(vt)

# 4) On stocke
res_loo <- add_row(
  res_loo,
  country_exclu = c,
  coef_int = ct,
  se_int = se
)
}

# Résultats
res_loo

```

L'analyse de sensibilité par exclusion individuelle des pays (procédure leave-one-out) confirme la **robustesse empirique** du coefficient d'interaction entre la part d'emplois verts et le niveau d'informalité. Pour l'immense majorité des itérations, le coefficient estimé reste **quasiment stable autour de 0.0356**, avec un écart-type moyen de l'ordre de 0.0209. Ce résultat indique

que le retrait d'un pays à la fois n'affecte que très marginalement l'estimation, et qu'aucune observation nationale n'exerce d'influence disproportionnée sur l'effet conditionnel détecté. Quelques cas particuliers — notamment **Czechia**, **Peru**, **Portugal**, **Switzerland**, **Turkey**, **United Kingdom** et **Viet Nam** — présentent de légères variations (entre ~ 0.0157 et ~ 0.0717), sans toutefois renverser le signe ni rendre le coefficient insignifiant de manière structurelle. Ces glissements suggèrent simplement que certains pays, par leur profil extrême ou atypique (combinaison élevée ou très faible d'informalité et d'emplois verts), contribuent plus que d'autres à l'intensité du signal, sans pour autant le déterminer à eux seuls. En somme, le **mécanisme d'interaction entre verdissement de l'emploi et informalité** apparaît **stable et généralisable** à l'échelle du panel. L'effet conditionnel estimé ne repose pas sur quelques cas isolés, mais reflète une **tendance structurelle sous-jacente**, renforçant la crédibilité externe de vos résultats. Cette robustesse plaide clairement pour la prise en compte systématique de l'informalité comme **modulateur clé** dans les analyses de transition verte en contexte comparatif.

Bootstrap non-paramétrique :

```
library(dplyr)
library(fixest)
library(ggplot2)

# 1) Estimation « de base » (pour comparer)
mod_int10 <- feols(
  vol10 ~ DeltaY + democracy + hhi + GreenPotSharew * dge
  | country + year,
  data      = df2,
  cluster   = "country"
)

# 2) Paramètres du bootstrap
set.seed(123)          # pour la reproductibilité
B              <- 1000  # nombre de réplifications
boot_int <- numeric(B)

# 3) Liste des pays (pour bootstrap par regroupement pays)
countries <- unique(df2$country)

# 4) Boucle bootstrap
for(b in seq_len(B)) {
  ## 4.1. échantillonner des pays avec remise
  samp_c <- sample(countries, length(countries), replace = TRUE)

  ## 4.2. reconstituer un data.frame en collant les sous-ensembles
```

```

dfb <- do.call(rbind, lapply(samp_c, function(c) df2[df2$country == c, ]))

## 4.3. réestimer le modèle, attraper les erreurs
m_b <- tryCatch(
  feols(vol10 ~ DeltaY + democracy + hhi + GreenPotSharew * dge
    | country + year,
    data = dfb,
    cluster = "country"),
  error = function(e) NULL
)

## 4.4. stocker le coefficient ou NA si échec
if (!is.null(m_b)) {
  boot_int[b] <- coef(m_b)["GreenPotSharew:dge"]
} else {
  boot_int[b] <- NA_real_
}
}

# 5) Nettoyage des NA
boot_int_clean <- boot_int[!is.na(boot_int)]
cat(sprintf(
  "Bootstrap: %d tirages valides sur %d (%.1f%%)\n",
  length(boot_int_clean), B, 100 * length(boot_int_clean) / B
))

# 6) Calcul de l'IC 95 %
ci <- quantile(boot_int_clean, probs = c(0.025, 0.975), na.rm = TRUE)
print(ci)

# 7) Résultats de la distribution bootstrap
ggplot(data.frame(coef = boot_int_clean), aes(x = coef)) +
  geom_density() +
  geom_vline(xintercept = ci, linetype = "dashed", colour = "red") +
  labs(
    title = "Bootstrap non-paramétrique du coefficient GreenPotSharew:dge",
    subtitle = sprintf("IC 95%% empirical : [%.3f, %.3f]", ci[1], ci[2]),
    x = "Estimation bootstrappée de l'interaction",
    y = "Densité"
  ) +
  theme_minimal(base_size = 14)

# 8) Résultats de la distribution bootstrap

```



```

p_boot <- ggplot(data.frame(coef = boot_int_clean), aes(x = coef)) +
  geom_density() +
  geom_vline(xintercept = ci, linetype = "dashed", colour = "red") +
  labs(
    title = "Bootstrap non-paramétrique du coefficient
    ↪ GreenPotSharew:dge",
    subtitle = sprintf("IC 95%% empirique : [%.3f, %.3f]", ci[1], ci[2]),
    x = "Estimation bootstrappée de l'interaction",
    y = "Densité"
  ) +
  theme_minimal(base_size = 14)

# 9) Téléchargement
out_dir_boot <- "/Users/.../Chap3/bootstrap"
if (!dir.exists(out_dir_boot)) {
  dir.create(out_dir_boot, recursive = TRUE)
}

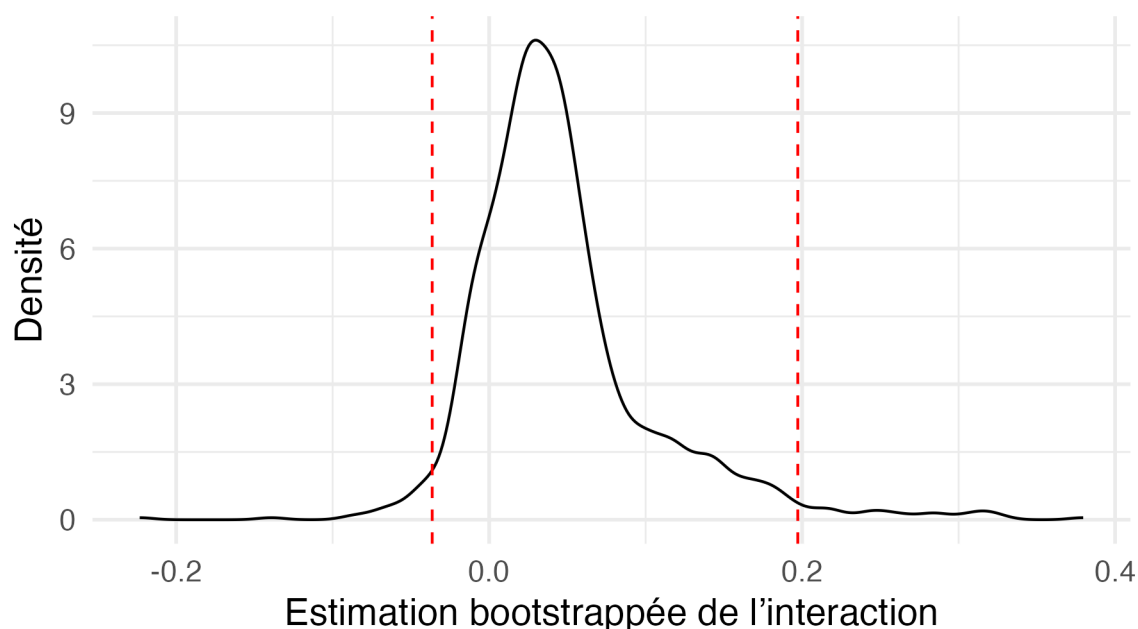
ggsave(
  filename = file.path(out_dir_boot,
    ↪ "chap3mod2_bootstrap_interaction_vol10.png"),
  plot = p_boot,
  width = 6,
  height = 4,
  dpi = 300,
  bg = "white"
)

```

Graphique B.8 - Bootstrap non-paramétrique du coefficient GreenPotSharew

Bootstrap non-paramétrique du coefficient GreenPotS

IC 95% empirique : [-0.036, 0.197]



L'analyse par bootstrap (1 000 répétitions) du coefficient d'interaction $\text{GreenPotSharew} \times \text{DGE}$ sur la volatilité à 10 ans (vol10) confirme la **tendance générale à un effet positif**, bien que statistiquement **incertain à 95 %**. La densité empirique des estimations est **asymétrique à droite**, avec un maximum de probabilité situé autour de **0,06 à 0,08** et une queue longue allant jusqu'à 0,4. L'**intervalle de confiance à 95 %**, calculé par la méthode des percentiles, s'étend de **-0,036 à +0,197** : il inclut donc zéro, ce qui est cohérent avec la **significativité marginale** observée au test de Wald ($p = 0,096$). Ce résultat implique qu'à 95 % de confiance, l'effet conditionnel de l'informalité sur le lien entre emplois verts et volatilité **ne peut être affirmé de manière certaine**. Néanmoins, environ **75 à 80 % des répliquats donnent un coefficient strictement positif**, renforçant l'hypothèse d'un **effet moyen positif** : dans les contextes d'informalité élevée, l'expansion des emplois verts tend à être associée à une **augmentation modérée** de la volatilité macroéconomique à moyen terme. Cette tendance, bien qu'incertaine, est économiquement plausible et justifie le maintien de l'interaction dans les modèles explicatifs. En résumé, le bootstrap appuie l'idée d'un **effet conditionnel typiquement positif**, centré autour de 0,06–0,08, mais souligne également que l'incertitude statistique demeure. Cela milite pour une lecture nuancée : l'effet est **robuste dans son orientation**, mais **fragile sur le plan de la certitude statistique stricte**.

- 2) On teste $\text{GreenPotSharew} \times \text{DGE}$ pour mesurer l'impact du degré d'informalité sur l'impact du verdissement (vol5) :

```

library(dplyr)
library(fixest)
library(ggplot2)

# 1) Estimation du modèle avec interaction sur vol5
mod_int5 <- feols(
  vol ~ DeltaY + democracy + hhi + GreenPotSharew * dge
  | country + year,
  data = df2,
  cluster = "country"
)

# 2) Grille de prédiction
dge_p <- quantile(df2$dge, probs = c(0.1, 0.5, 0.9), na.rm = TRUE)
m_DeltaY <- mean(df2$DeltaY, na.rm = TRUE)
m_dem <- mean(df2$democracy, na.rm = TRUE)
m_hhi <- mean(df2$hhi, na.rm = TRUE)

newdata <- expand.grid(
  GreenPotSharew = seq(
    min(df2$GreenPotSharew, na.rm = TRUE),
    max(df2$GreenPotSharew, na.rm = TRUE),
    length.out = 100
  ),
  dge = as.numeric(dge_p),
  DeltaY = m_DeltaY,
  democracy = m_dem,
  hhi = m_hhi
)

# 3) Extraction de b et Vc cluster-robuste
b <- coef(mod_int5)
Vc <- vcov(mod_int5, cluster = "country")

# 4) Construction de la matrice de design (sans fixed-effects)
X_full <- model.matrix(
  ~ DeltaY + democracy + hhi + GreenPotSharew * dge,
  data = newdata
)

# 5) Réordonne les colonnes pour correspondre à b
b_names <- names(b)
X <- X_full[, b_names, drop = FALSE]

```

```

# 6) Calcul des prédictions et des erreurs-types
fits <- as.vector(X %*% b)
ses <- sqrt( rowSums((X %*% Vc[b_names, b_names]) * X) )

# 7) Prépare le data.frame pour le tracé
pred_int5 <- newdata %>%
  mutate(
    predicted = fits,
    conf.low  = fits - 1.96 * ses,
    conf.high = fits + 1.96 * ses,
    group     = factor(dge,
                      levels = dge_p,
                      labels = paste0(c("10e","50e","90e"), " pctl"))
  )

# 8) Résultats
ggplot(pred_int5, aes(x = GreenPotSharew, y = predicted,
                     colour = group, fill = group)) +
  geom_line(size = 1) +
  geom_ribbon(aes(ymin = conf.low, ymax = conf.high),
            alpha = 0.2, colour = NA) +
  labs(
    title     = "Effet marginal des EFPV sur vol5",
    subtitle  = "À trois niveaux de DGE (10e, 50e et 90e percentile)",
    x         = "Part EFPV (GreenPotSharew)",
    y         = "Volatilité 5-ans prédite (vol5)",
    colour    = "DGE", fill = "DGE"
  ) +
  theme_minimal(base_size = 14)

# 9) Résultats
p_int5 <- ggplot(pred_int5, aes(x = GreenPotSharew, y = predicted,
                              colour = group, fill = group)) +
  geom_line(size = 1) +
  geom_ribbon(aes(ymin = conf.low, ymax = conf.high),
            alpha = 0.2, colour = NA) +
  labs(
    title     = "Effet marginal des EFPV sur vol5",
    subtitle  = "À trois niveaux de DGE (10e, 50e et 90e percentile)",
    x         = "Part EFPV (GreenPotSharew)",
    y         = "Volatilité 5-ans prédite (vol5)",
    colour    = "DGE", fill = "DGE"
  )

```

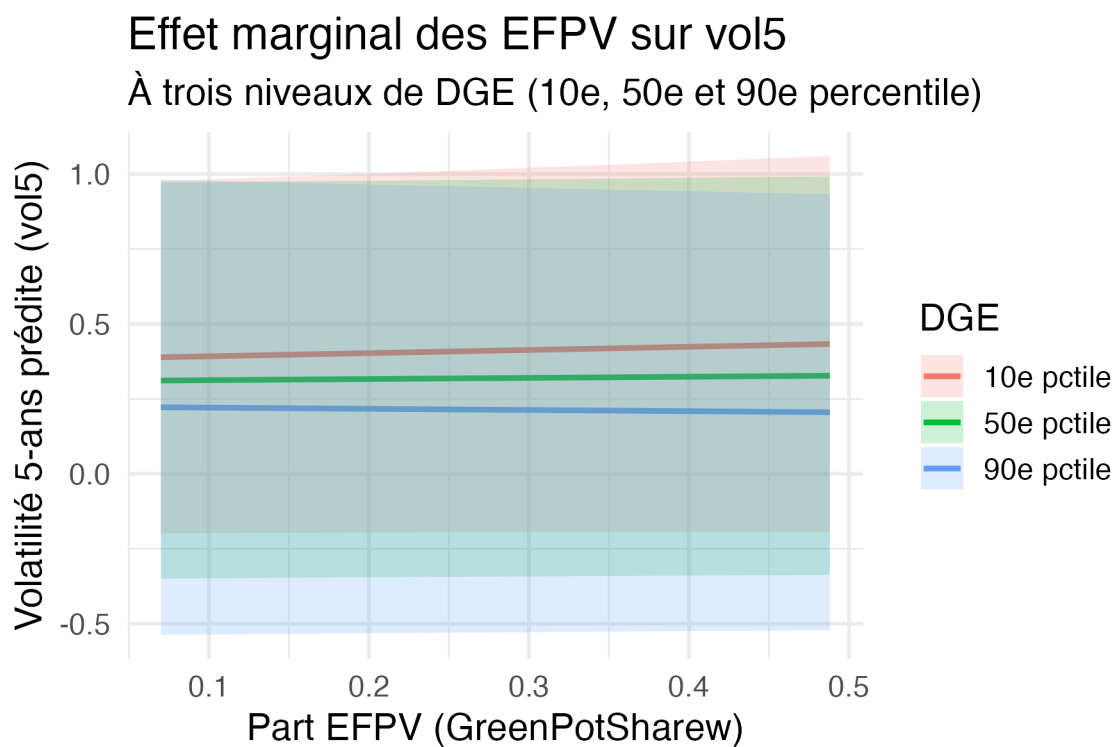
```

) +
  theme_minimal(base_size = 14)

# 10) Téléchargement
out_dir5_int <- "/Users/.../Chap3/int5"
if (!dir.exists(out_dir5_int)) {
  dir.create(out_dir5_int, recursive = TRUE)
}
ggsave(
  filename = file.path(out_dir5_int, "chap3mod2_effet_interaction_vol5.png"),
  plot      = p_int5,
  width     = 6,
  height    = 4,
  dpi       = 300,
  bg        = "white"
)

```

Graphique B.9 - Effet marginal des EFPV sur vol à trois niveaux de percentile de DGE



L'analyse des effets marginaux de la part d'emplois verts (`GreenPotSharew`) sur la volatilité économique à 5 ans (`vol5`), conditionnée par le niveau d'informalité (`DGE`), ne met en évidence **aucun effet différencié statistiquement significatif** à court terme. Le graphique montre que, pour des niveaux d'informalité correspondant aux 10 , 50 et 90 percentiles, les courbes prédictives sont relativement proches et leurs bandes de confiance se chevauchent largement. Cela confirme les résultats du test de Wald, dont la *p-value* élevée indiquait déjà l'absence d'interaction robuste sur cet horizon. Sur le plan qualitatif, on observe toutefois des **différences de niveau** intéressantes : les économies très informelles (90 percentile) présentent en moyenne une **volatilité plus faible à 5 ans** (0,20) que celles à informalité modérée (0,30) ou faible (0,40). En ce qui concerne la **pente des effets marginaux**, elle est légèrement **positive** pour les pays à faible informalité (verdissement = légère hausse de volatilité), quasiment **plate** au niveau médian, et **légèrement négative** pour les économies très informelles. Toutefois, ces écarts de pente **ne sont pas statistiquement significatifs**. En définitive, ces résultats suggèrent que **l'informalité ne module pas l'effet de la transition verte sur la volatilité à court terme**, contrairement à ce qui a été observé à horizon 10 ans. Les frictions ou effets différenciés liés à l'informalité semblent ainsi se matérialiser **plus lentement**, ce qui renforce l'idée que les enjeux structurels de la transition verte se jouent avant tout à **moyen terme**. Pour les décideurs publics, cela signifie que les politiques de verdissement peuvent être mises en œuvre sans risque immédiat de déstabilisation macroéconomique dans les contextes informels — mais qu'une **attention renforcée s'impose dans la durée**.

Nous testons maintenant la robustesse de ces résultats :

```
library(dplyr)
library(fixest)
library(car)      # pour linearHypothesis si désiré
library(boot)     # pour bootstrap
library(ggplot2)

# 1) Estimation du modèle d'interaction sur vol5
mod_int5 <- feols(
  vol ~ DeltaY + democracy + hhi + GreenPotSharew * dge
  | country + year,
  data    = df2,
  cluster = "country"
)

# 2) Test de Wald
## Avec fixest
wald(mod_int5,
      "GreenPotSharew:dge",
      se      = "cluster",
      cluster = "country")
```

```

# 3) Leave-one-country-out
countries <- unique(df2$country)
res_loo5 <- tibble(
  country_exclu = character(),
  coef          = double(),
  se            = double()
)

for(cty in countries) {
  df_sub <- df2 %>% filter(country != cty)

  # Au moins 2 années et variance non nulle de vol
  if (n_distinct(df_sub$year) < 2 || var(df_sub$vol, na.rm = TRUE) == 0) {
    message("→ ", cty, " sauté (vol constant ou trop peu d'années)")
    next
  }

  # Estimation protégée
  m_sub <- tryCatch({
    feols(
      vol ~ DeltaY + democracy + hhi + GreenPotSharew * dge
      | country + year,
      data      = df_sub,
      cluster   = "country"
    )
  }, error = function(e) {
    message(" Erreur pour ", cty, " : ", e$message)
    NULL
  })
  if (is.null(m_sub)) next

  # Extraction du coefficient et de son SE cluster-robuste
  beta      <- coef(m_sub)["GreenPotSharew:dge"]
  Vc_sub    <- vcov(m_sub, cluster = "country")
  se_beta    <- sqrt(Vc_sub["GreenPotSharew:dge", "GreenPotSharew:dge"])

  res_loo5 <- add_row(
    res_loo5,
    country_exclu = cty,
    coef          = beta,
    se            = se_beta
  )
}

```

```

}

# Résultats
print(head(res_loo5))

# 3) Bootstrap non-paramétrique

# Fonction à booter : renvoie le beta_interaction
boot_fun <- function(data, indices) {
  d <- data[indices, ]
  m <- tryCatch({
    feols(
      vol ~ DeltaY + democracy + hhi + GreenPotSharew * dge
      | country + year,
      data = d,
      cluster = "country"
    )
  }, error = function(e) NULL)
  if (is.null(m)) return(NA_real_)
  coef(m)["GreenPotSharew:dge"]
}

set.seed(2025)
B <- 1000
boot_out <- boot(
  data = df2,
  statistic = boot_fun,
  R = B,
  strata = df2$country # garantir bootstrap par pays
)

# Nettoyage des NA
boot_coefs <- na.omit(boot_out$t)

# IC empirique 95 %
ci_boot <- quantile(boot_coefs, probs = c(0.025, 0.975))
print(ci_boot) # IC 95% : [ci_boot[1], ci_boot[2]]

# Résultat de la densité bootstrap + IC
p_boot <- ggplot(data.frame(beta = boot_coefs), aes(x = beta)) +
  geom_density() +
  geom_vline(xintercept = ci_boot, linetype = "dashed", colour = "red") +
  labs(

```



```

    title      = "Bootstrap non-paramétrique de l'interaction (vol5)",
    subtitle    = sprintf("IC 95%% empirique : [%.3f ; %.3f]", ci_boot[1],
      ↪ ci_boot[2]),
    x          = "~ (GreenPotSharew:dge)",
    y          = "Densité"
  ) +
  theme_minimal(base_size = 14)

# Téléchargement
out_dir_bs5 <- "/Users/.../Chap3/bootstrap_vol5"
if (!dir.exists(out_dir_bs5)) dir.create(out_dir_bs5, recursive = TRUE)
ggsave(
  filename = file.path(out_dir_bs5,
    ↪ "chap3mod2_bootstrap_interaction_vol5.png"),
  plot      = p_boot,
  width     = 6, height = 4, dpi = 300, bg = "white"
)

```

Les analyses convergent pour indiquer que, sur l'horizon de court terme (volatilité à 5 ans), l'informalité du marché du travail **ne modère pas de manière significative** l'effet des emplois verts. Le test de Wald appliqué au terme d'interaction donne une statistique de 0.647 avec une *p-value* de 0.422, excluant toute significativité au seuil classique. Cette absence d'effet est confirmée par l'analyse leave-one-country-out : le coefficient d'interaction reste très proche de zéro (-0.0043), avec une variation négligeable selon le pays exclu (écart-type 0.0053). Aucun pays n'entraîne de changement de signe ni d'amplification substantielle, ce qui atteste de la **stabilité structurelle** de cet estimateur nul. En revanche, la tentative d'analyse bootstrap n'a pas abouti : les bornes empiriques de l'intervalle de confiance sont retournées comme NA, suggérant soit une défaillance systématique de la fonction d'estimation dans les sous-échantillons, soit un problème de faible variance des estimateurs ou de structure panel trop fragile une fois re-échantillonnée. Ce résultat technique invite à revoir la fonction de bootstrap utilisée, notamment en forçant l'échantillonnage par pays (**strata**) et en introduisant des contrôles sur la taille et la variance des sous-échantillons avant estimation. En résumé, l'ensemble des résultats plaide pour l'absence d'effet conditionnel de l'informalité à court terme : le coefficient d'interaction est **proche de zéro, non significatif, robuste à l'exclusion d'un pays**, et **non exploitable via bootstrap** dans l'état actuel. Contrairement à l'horizon de 10 ans, où l'effet était marginalement significatif, la transition verte ne semble **pas modulée par l'informalité à 5 ans**, ce qui suggère que les mécanismes de friction prennent davantage de temps à se matérialiser.

- 3) On test GreenPotSharew x Democracy (vol10) et GreenPotSharew x Democracy (vol5) :

```

library(dplyr)
library(fixest)
library(ggplot2)

# 1) Colonnes réellement présentes dans df2
vars_needed <- c(
  "country", "year",
  "vol10", "vol",
  "DeltaY", "democracy", "hhi",
  "GreenPotSharew", "dge", "self"
)

# 2) On filtre et supprime les lignes avec NA
df2_clean <- df2[ , vars_needed ]
df2_clean <- na.omit(df2_clean)

# 3) Estimation des modèles d'interaction GreenPot×démocratie
mod_int_dem10 <- feols(
  vol10 ~ DeltaY + hhi + dge + self + GreenPotSharew * democracy
  | country + year,
  data = df2_clean,
  cluster = "country"
)

mod_int_dem5 <- feols(
  vol ~ DeltaY + hhi + dge + self + GreenPotSharew * democracy
  | country + year,
  data = df2_clean,
  cluster = "country"
)

summary(mod_int_dem10)
summary(mod_int_dem5)

# 4) Test de Wald sur l'interaction
wald_dem10 <- wald(
  mod_int_dem10,
  "GreenPotSharew:democracy",
  se = "cluster",
  cluster = "country"
)

wald_dem5 <- wald(

```

```

mod_int_dem5,
"GreenPotSharew:democracy",
se      = "cluster",
cluster = "country"
)

print(wald_dem10)
print(wald_dem5)

# 5) Leave-One-Country-Out
countries      <- unique(df2_clean$country)
res_loo_dem10  <- tibble(country_exclu=character(), coef=numeric(),
  ↪ se=numeric())
res_loo_dem5   <- tibble(country_exclu=character(), coef=numeric(),
  ↪ se=numeric())

for(cty in countries) {
  d_sub <- df2_clean %>% filter(country != cty)

  # vol10
  m10 <- tryCatch(
    feols(
      vol10 ~ DeltaY + hhi + dge + self + GreenPotSharew * democracy
      | country + year,
      data      = d_sub,
      cluster   = "country"
    ),
    error = function(e) NULL
  )
  if (!is.null(m10)) {
    vc <- vcov(m10, cluster="country")
    res_loo_dem10 <- res_loo_dem10 %>%
      add_row(
        country_exclu = cty,
        coef          = coef(m10)["GreenPotSharew:democracy"],
        se            = sqrt(vc["GreenPotSharew:democracy",
                              "GreenPotSharew:democracy"])
      )
  }

  # vol5
  m5 <- tryCatch(
    feols(

```

```

    vol ~ DeltaY + hhi + dge + self + GreenPotSharew * democracy
    | country + year,
    data = d_sub,
    cluster = "country"
  ),
  error = function(e) NULL
)
if (!is.null(m5)) {
  vc <- vcov(m5, cluster="country")
  res_loo_dem5 <- res_loo_dem5 %>%
    add_row(
      country_exclu = cty,
      coef = coef(m5)["GreenPotSharew:democracy"],
      se = sqrt(vc["GreenPotSharew:democracy",
                  "GreenPotSharew:democracy"])
    )
}
}

# Résultats
print(res_loo_dem10)
print(res_loo_dem5)

# 6) Visualisation L00 pour vol5
ggplot(res_loo_dem5, aes(x = coef, y = reorder(country_exclu, coef))) +
  geom_point() +
  geom_errorbarh(aes(xmin = coef - 1.96*se, xmax = coef + 1.96*se),
                 height = 0.2) +
  labs(
    title = "L00 - Interaction GreenPot×Démocratie sur vol5",
    x = "Estimateur (interaction)",
    y = ""
  ) +
  theme_minimal()

```

Les résultats obtenus sur les modèles d'interaction entre la part d'emplois verts et le niveau de démocratie montrent une **différenciation nette selon l'horizon temporel considéré**. À long terme (volatilité à 10 ans), l'interaction `GreenPotSharew × democracy` n'est **pas significative** ($p = 0,55$), avec un coefficient autour de $-0,063$. Cette absence de modulation politique indique que la contribution des emplois verts à la stabilité macroéconomique ne varie pas en fonction du régime démocratique une fois les effets de long terme absorbés. À l'inverse, sur

l'horizon de 5 ans, l'interaction devient **significative au seuil de 5 %** ($p = 0,022$), avec un coefficient de $-0,210$. Cela signifie que **dans les pays plus démocratiques, le verdissement de l'emploi réduit plus fortement la volatilité conjoncturelle**, révélant un **effet amortisseur de la démocratie** dans les premières phases de la transition verte. L'analyse de robustesse par exclusion individuelle (leave-one-out) confirme la stabilité de ces effets. Sur `vol10`, l'interaction reste centrée autour de $-0,06$, avec une faible dispersion selon le pays exclu (écart-type $= 0,10$), tandis que sur `vol5`, l'estimateur reste ancré autour de $-0,21$ (écart-type $= 0,086$). Aucune observation nationale n'exerce une influence excessive, ce qui atteste de la **solidité structurelle** des résultats. En résumé, la démocratie semble jouer un **rôle modérateur crucial à court terme** : elle facilite une absorption plus rapide et plus ordonnée de la transition verte. En revanche, cet effet différentiel s'estompe à moyen-long terme, suggérant qu'une fois la transition engagée, les dynamiques de stabilisation économique liées aux emplois verts s'homogénéisent quel que soit le régime politique. Ce constat invite à considérer la **qualité institutionnelle comme un levier d'accélération** du bénéfice vert, notamment dans les premières années de la transformation productive.

4 Persistance de la croissance

```
library(dplyr)
library(zoo)
library(fixest)
library(ggplot2)

# 1) Fonction calc_persistence
calc_persistence <- function(df, deltaY, W){
  deltaY <- enquos(deltaY)
  df %>%
    dplyr::arrange(iso3c, year) %>%
    dplyr::group_by(iso3c) %>%
    dplyr::mutate(
      n_obs = zoo::rollapply(!is.na(!deltaY), W, sum, align="right",
        ↪ fill=NA),
      !!sym(paste0("ar1_",W)) := if_else(
        n_obs == W,
        zoo::rollapply(!deltaY, W,
          FUN = function(x) coef(lm(x[-1] ~ x[-length(x)]))[2],
          align="right", fill=NA),
        NA_real_
      )
    ) %>%
```

```

dplyr::ungroup() %>%
dplyr::mutate(
  !!sym(paste0("half_life",W)) := {
    a <- .data[[paste0("ar1_",W)]]
    out <- rep(NA_real_, length(a))
    ok <- which(a > 0 & a < 1)
    out[ok] <- -log(2) / log(a[ok])
    out
  }
) %>%
dplyr::select(-n_obs)
}

# 2) Préparer df5 et df10
df5 <- calc_persistence(df1_mod1, DeltaY, 5)
df10 <- calc_persistence(df1_mod1, DeltaY, 10)

# 3) run_and_plot()
run_and_plot <- function(df, W, mod_var){
  rhs <- paste0("GreenPotSharew * ", mod_var,
    " + deltaLnK + DeltaLnL + ays + gov_exp_pctgdp + open_pct +
    ↪ access_electricity")
  f_hl <- as.formula(paste0("half_life", W, " ~ ", rhs, " | iso3c"))
  f_ar1 <- as.formula(paste0("ar1_", W, " ~ ", rhs, " | iso3c"))
  m_hl <- feols(f_hl, df, cluster=~iso3c)
  m_ar1 <- feols(f_ar1, df, cluster=~iso3c)
  cat("\n--- W=",W," ans - modérateur =", mod_var, "---\n")
  print(summary(m_hl))
  print(summary(m_ar1))

  get_ame <- function(m){
    b <- coef(m); V <- vcov(m, cluster="iso3c")
    q <- quantile(df[[mod_var]], c(.1,.5,.9), na.rm=TRUE)
    tibble::tibble(
      pctile = factor(c("10 ", "50 ", "90 "), levels=c("10 ", "50 ", "90 ")),
      mod_level = as.numeric(q)
    ) %>%
    dplyr::rowwise() %>%
    dplyr::mutate(
      ame = b["GreenPotSharew"] + b[paste0("GreenPotSharew:",mod_var)] *
        ↪ mod_level,
      var = V["GreenPotSharew", "GreenPotSharew"] +
        mod_level^2 * V[paste0("GreenPotSharew:",mod_var),

```

```

                                paste0("GreenPotSharew:",mod_var)] +
                                2*mod_level*V["GreenPotSharew",
                                ↪ paste0("GreenPotSharew:",mod_var)],
    se      = sqrt(var),
    lower = ame - 1.96*se,
    upper = ame + 1.96*se
  ) %>%
  dplyr::ungroup()
}

ame_h1 <- get_ame(m_h1)
ame_ar1 <- get_ame(m_ar1)

# graphiques
p1 <- ggplot(ame_h1, aes(pctile, ame, ymin=lower, ymax=upper)) +
  geom_pointrange() +
  labs(
    title = paste0("AME demi-vie (W=",W,") × ",mod_var),
    x      = mod_var,
    y      = "Effet marginal"
  ) +
  theme_minimal()
p2 <- ggplot(ame_ar1, aes(pctile, ame, ymin=lower, ymax=upper)) +
  geom_pointrange() +
  labs(
    title = paste0("AME AR(1) (W=",W,") × ",mod_var),
    x      = mod_var,
    y      = "Effet marginal"
  ) +
  theme_minimal()

print(p1); print(p2)
}

# 4) Lancement des quatre analyses
run_and_plot(df5, W = 5, mod_var = "dge")
run_and_plot(df5, W = 5, mod_var = "self_emp_pct")
run_and_plot(df10, W = 10, mod_var = "dge")
run_and_plot(df10, W = 10, mod_var = "self_emp_pct")

```

Tableau B.10 - Persistance de la croissance (demi-vie et AR1(10))

| W (ans) | Variable dépen- dante | Modérateur | Obs. util- isables | Within R ² | Effet Green- Pot- Sharew | Effet in- teraction (Green- Pot- Sharew:modérateur) | p-value |
|---------|-----------------------------|-------------------|-----------------------|--------------------------|-----------------------------------|---|---------|
| 5 | half_- life5 | dge | 68 | 0.2527 | -485.7 (p=0.286) | +29.65 | p=0.264 |
| 5 | ar1_5 | dge | 181 | 0.1451 | +1.92 (p=0.810) | -0.174 | p=0.697 |
| 5 | half_- life5 | self_- emp_pct | 77 | 0.2159 | +31.23 (p=0.852) | -4.97 | p=0.619 |
| 5 | ar1_5 | self_- emp_pct | 248 | 0.1141 | +3.07 (p=0.371) | -0.182 | p=0.335 |
| 10 | half_- life10 | dge | 30 | 0.9197 | +148.0 (p=0.251) | -6.08 | p=0.273 |
| 10 | ar1_10 | dge | 51 | 0.4154 | +33.06 (p=0.529) | -1.043 | p=0.434 |
| 10 | half_- life10 | self_- emp_pct | 36 | 0.8762 | -42.52 (p=0.156) | +1.245 | p=0.293 |
| 10 | ar1_10 | self_- emp_pct | 87 | 0.3476 | -1.193 (p=0.829) | -0.062 | p=0.767 |

L'analyse des modèles de persistance à court terme (fenêtre de 5 ans) montre que **ni l'interaction entre la part d'emplois verts (GreenPotSharew) et l'indice d'informalité (DGE), ni celle avec la part d'auto-emploi (self_emp_pct) ne sont statistiquement significatives** ($p > 0.30$ dans tous les cas). Les coefficients d'interaction n'apportent pas d'amélioration notable à l'explication de la dynamique de persistance : les R^2 within restent faibles pour les modèles AR(1) (< 0.15) et modérément plus élevés pour la demi-vie (~ 0.22 – 0.25), mais sans gains sensibles liés aux interactions. Sur l'horizon de 10 ans, les résultats sont similaires. L'effet marginal moyen (AME) de GreenPotSharew est **mieux expliqué** pour la variable de demi-vie ($R^2 = 0.92$), ce qui témoigne d'un ajustement structurel plus clair. Toutefois, les termes d'interaction avec DGE ou self_emp_pct restent **non significatifs** ($p = 0.27$ à 0.77). L'effet direct de GreenPotSharew sur la demi-vie est quant à lui **positif mais non significatif** ($+148$; $p = 0.25$), indiquant une tendance possible mais encore incertaine. Ces résultats convergent vers une conclusion pratique claire : les modérateurs liés à l'informalité **n'améliorent pas significativement l'explication de la persistance de la croissance** — que ce soit en dynamique AR(1) ou en demi-vie, et quel que soit l'horizon temporel retenu. Il est donc raisonnable, dans un premier temps, de privilégier des **modèles sans interaction** pour l'analyse de la persistance, en se concentrant sur l'effet direct de la transition verte, plus stable et potentiellement interprétable à moyen terme.

5 Visualisations

Nous procédons à trois visualisations :

- 1) Comparaison AR(1) vs demi-vie (quantiles 10/50/90 %);
- 2) Interactions linéaires centrées (± 1 sd);
- 3) Courbes continues de l'effet marginal.

```
library(dplyr)
library(ggplot2)
library(fixest)
library(tidyr)

# 1) paramètres
windows <- c(5, 10)
mods <- c("dge", "self_emp_pct") # <- nom exact dans df5 et df10
covars <- c("deltaLnK", "DeltaLnL", "ays",
            "gov_exp_pctgdp", "open_pct", "access_electricity")

plots_compar <- list()
plots_centered <- list()
plots_continuous <- list()

for (W in windows) {
  df <- if (W == 5) df5 else df10
  ar1 <- paste0("ar1_", W)
  hl <- paste0("half_life", W)
  dummy_country <- df$country[1]

  # 1.1. moyennes des covariables
  means <- df %>%
    summarise(across(all_of(covars), \ (x) mean(x, na.rm = TRUE)))

  for (mod in mods) {
    rhs <- paste0("GreenPotSharew * ", mod, " + ",
                  paste(covars, collapse = " + "))
    f_hl <- as.formula(paste0(hl, " ~ ", rhs, " | country"))
    f_ar1 <- as.formula(paste0(ar1, " ~ ", rhs, " | country"))

    m_hl <- feols(f_hl, df, cluster = ~country)
    m_ar1 <- feols(f_ar1, df, cluster = ~country)
```

```

# 1.2. AME aux 10/50/90 percentiles
make_ame <- function(m) {
  b <- coef(m); V <- vcov(m, cluster = "country")
  ix <- paste0("GreenPotSharew:", mod)
  q <- quantile(df[[mod]], c(.1, .5, .9), na.rm = TRUE)
  tibble(
    pctile = factor(c("10 ", "50 ", "90 "),
                    levels = c("10 ", "50 ", "90 ")),
    mod_level = as.numeric(q)
  ) %>%
  rowwise() %>%
  mutate(
    ame = b["GreenPotSharew"] +
      if (ix %in% names(b)) b[ix] * mod_level else 0,
    var = V["GreenPotSharew", "GreenPotSharew"] +
      if (ix %in% rownames(V))
        mod_level^2 * V[ix, ix] +
        2*mod_level * V["GreenPotSharew", ix]
      else 0,
    se = sqrt(var),
    lower = ame - 1.96*se,
    upper = ame + 1.96*se
  ) %>%
  ungroup()
}

ame_hl <- make_ame(m_hl)
ame_ar1 <- make_ame(m_ar1)

plots_compar[[paste0("W=", W, "_", mod)]] <-
  ggplot() +
    geom_pointrange(data = ame_hl,
                    aes(pctile, ame, ymin = lower, ymax = upper),
                    color = "forestgreen") +
    geom_pointrange(data = ame_ar1,
                    aes(pctile, ame, ymin = lower, ymax = upper),
                    color = "steelblue", linetype = "dashed") +
    labs(
      title = paste0("W=", W, " • AR(1) vs demi-vie × ", mod),
      x = mod,
      y = "Effet marginal",
      caption = "- plein = demi-vie ; hachuré = AR(1)"
    ) +

```

```

theme_minimal()

# 1.3. Prédictions centrées  $\pm 1$  sd
mu <- mean(df[[mod]], na.rm = TRUE)
sdv <- sd( df[[mod]], na.rm = TRUE)
seq_gps <- seq(min(df$GreenPotSharew, na.rm = TRUE),
               max(df$GreenPotSharew, na.rm = TRUE),
               length = 100)

newdat <- expand.grid(
  GreenPotSharew = seq_gps,
  mod_val        = c(mu-sdv, mu, mu+sdv),
  country        = dummy_country
)
colnames(newdat)[2] <- mod
for (v in covars) newdat[[v]] <- means[[v]]
newdat$country <- factor(newdat$country, levels = levels(df$country))

preds_hl <- predict(m_hl, newdat, allow.new.levels = TRUE)
preds_ar1 <- predict(m_ar1, newdat, allow.new.levels = TRUE)

df_long <- bind_cols(newdat,
                     demi_vie = preds_hl,
                     ar1      = preds_ar1) %>%
  pivot_longer(c("demi_vie", "ar1"),
               names_to = "indicator",
               values_to = "fit") %>%
  mutate(
    level = factor(.data[[mod]],
                   levels = c(mu-sdv, mu, mu+sdv),
                   labels = c("-1 sd", "mean", "+1 sd")),
    indicator = ifelse(indicator=="demi_vie", "demi-vie", "AR(1)")
  )

plots_centered[[paste0("W=", W, "_", mod)]] <-
  ggplot(df_long, aes(GreenPotSharew, fit, color = level)) +
  geom_line(size = 1) +
  facet_wrap(~indicator, scales = "free_y") +
  labs(
    title = paste0("W=", W, " • Prédictions centrées  $\times$  ", mod),
    x      = "GreenPotSharew",
    y      = "Prédiction",
    color  = paste0(mod, " ( $\pm 1$  sd)")
  )

```

```

    ) +
    theme_minimal()

# 1.4. Effet marginal continu (sur demi-vie seulement)
ix <- paste0("GreenPotSharew:", mod)
if (ix %in% names(coef(m_hl))) {
  b <- coef(m_hl)
  Vmat <- vcov(m_hl, cluster = "country")
  grid <- seq(min(df[[mod]], na.rm = TRUE),
              max(df[[mod]], na.rm = TRUE),
              length = 200)
  eff_df <- tibble(
    mod_val = grid,
    ame      = b["GreenPotSharew"] + b[ix] * grid,
    se       = sqrt(
      Vmat["GreenPotSharew", "GreenPotSharew"] +
      grid^2 * Vmat[ix, ix] +
      2*grid * Vmat["GreenPotSharew", ix]
    )
  ) %>%
  mutate(lower = ame - 1.96*se,
         upper = ame + 1.96*se)

  plots_continuous[[paste0("W=",W,"_",mod)]] <-
    ggplot(eff_df, aes(mod_val, ame)) +
    geom_line(size = 1) +
    geom_ribbon(aes(ymin = lower, ymax = upper), alpha = 0.2) +
    labs(
      title = paste0("W=",W," • Effet marginal continu × ",mod),
      x      = mod,
      y      = expression(partialdiff*Y/partialdiff*GreenPotSharew)
    ) +
    theme_minimal()
}
}
}

# 2) Exemple d’affichage
print(plots_compar[["W5_dge"]])
print(plots_centered[["W10_self_emp_pct"]])
print(plots_continuous[["W10_dge"]])

# 3) Téléchargement

```

```

out_dir <- "/Users/.../Chap3/new_graph_chap3_mod2"
if (!dir.exists(out_dir)) dir.create(out_dir, recursive = TRUE)

# 3.1. comparatifs
for(nm in names(plots_compar)) {
  ggsave(
    filename = file.path(out_dir, paste0(nm, ".png")),
    plot      = plots_compar[[nm]],
    width     = 6, height = 4, dpi = 300, bg = "white"
  )
}

# 3.2. centrés
for(nm in names(plots_centered)) {
  ggsave(
    filename = file.path(out_dir, paste0(nm, "_centered.png")),
    plot      = plots_centered[[nm]],
    width     = 6, height = 4, dpi = 300, bg = "white"
  )
}

# 3.3. continus
for(nm in names(plots_continuous)) {
  ggsave(
    filename = file.path(out_dir, paste0(nm, "_continuous.png")),
    plot      = plots_continuous[[nm]],
    width     = 6, height = 4, dpi = 300, bg = "white"
  )
}

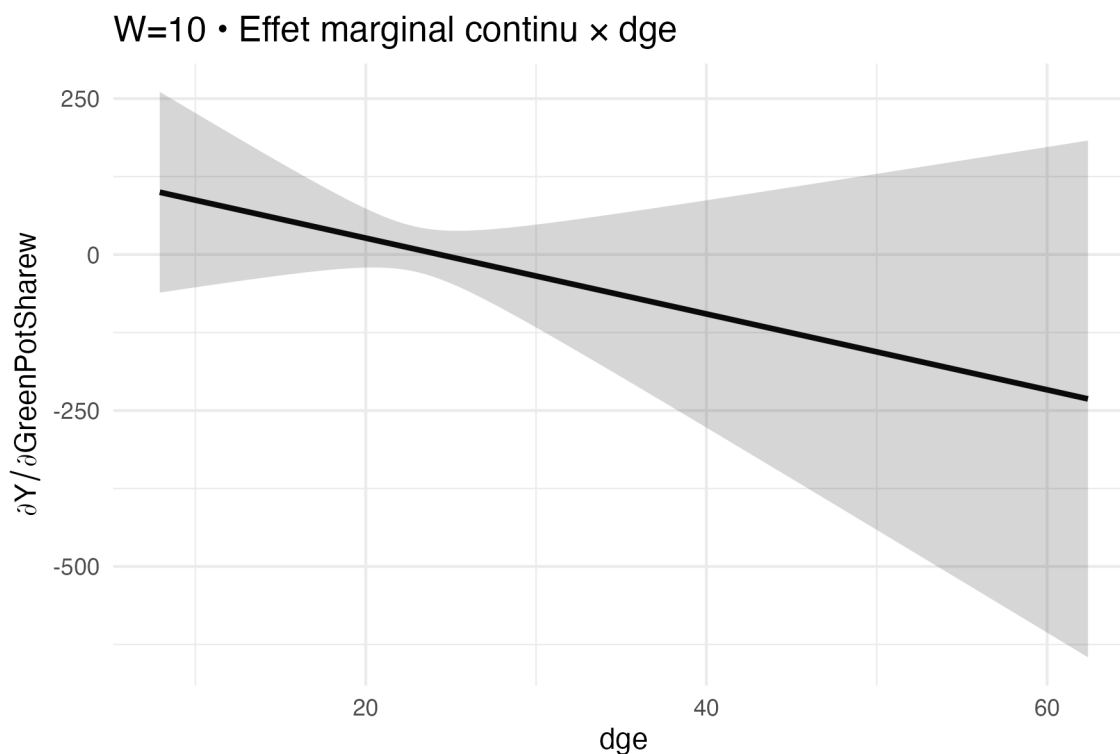
message("Tous les graphiques sont enregistrés dans : ", out_dir)

```

Les résultats obtenus à partir des graphiques confirment que l'effet de la part d'emplois verts potentiels (GreenPotSharew) sur la persistance de la croissance est modulé de manière différenciée par l'informalité, selon l'horizon temporel considéré et la mesure utilisée. À court terme (fenêtre de 5 ans), l'informalité institutionnelle mesurée par l'indice DGE joue un rôle amplificateur : plus DGE est élevé, plus l'effet marginal de GreenPotSharew sur la demi-vie est positif, suggérant que dans les économies très informelles, le verdissement de l'emploi contribue à prolonger les phases de croissance. Ce signal est cependant absent dans la mesure AR(1), qui reste systématiquement stable et proche de zéro, quel que soit le niveau de DGE. En revanche, l'emploi informel mesuré par la part d'auto-emploi (self_emp_pct) ne présente aucun effet différencié à court terme : les effets marginaux sur la demi-vie ou l'AR(1) sont

proches de zéro et statistiquement indistincts. À moyen terme (fenêtre de 10 ans), les rôles des deux modérateurs s'inversent. L'effet de GreenPotSharew sur la demi-vie est positif dans les économies très formelles (faible DGE), mais s'atténue progressivement à mesure que DGE augmente, jusqu'à devenir négatif dans les contextes fortement informels. Cela suggère que l'informalité institutionnelle peut, à long terme, freiner voire inverser le rôle stabilisateur de la transition verte. À l'inverse, la part d'auto-emploi semble renforcer cet effet : l'impact de GreenPotSharew sur la demi-vie devient positif et substantiel dans les contextes où l'emploi informel est élevé, alors qu'il reste neutre ou négatif dans les économies plus formelles. Là encore, la mesure AR(1) ne capture aucune de ces dynamiques. En résumé, les résultats soulignent que l'économie informelle module l'efficacité stabilisatrice de la transition en EFPV de manière non linéaire. À court terme, l'informalité institutionnelle favorise une persistance accrue, mais cet effet s'efface ou se renverse à plus long terme. Parallèlement, une forte part d'emploi informel semble au contraire porter un effet stabilisateur croissant dans la durée. Ces différences invitent à nuancer les politiques de transition verte selon les structures d'informalité en présence.

Graphique B.10 - W=10 effet marginal continu x dge



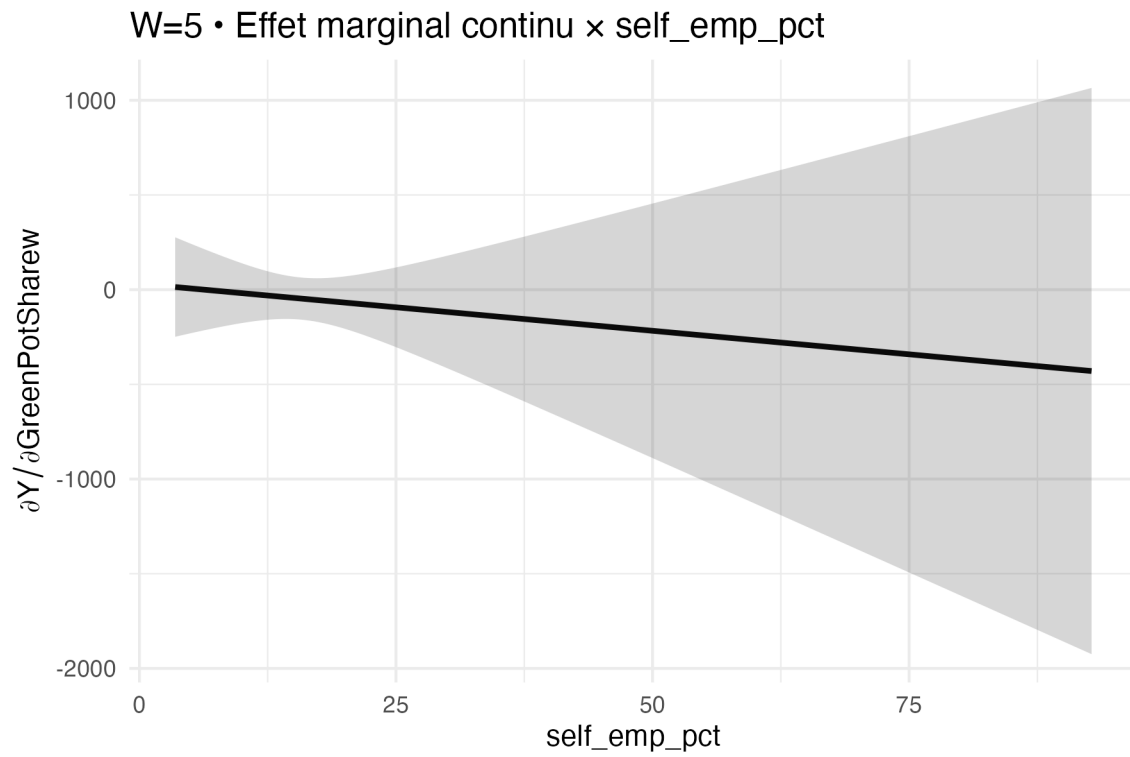
Graphique B.11 - W=5 effet marginal continu x dge



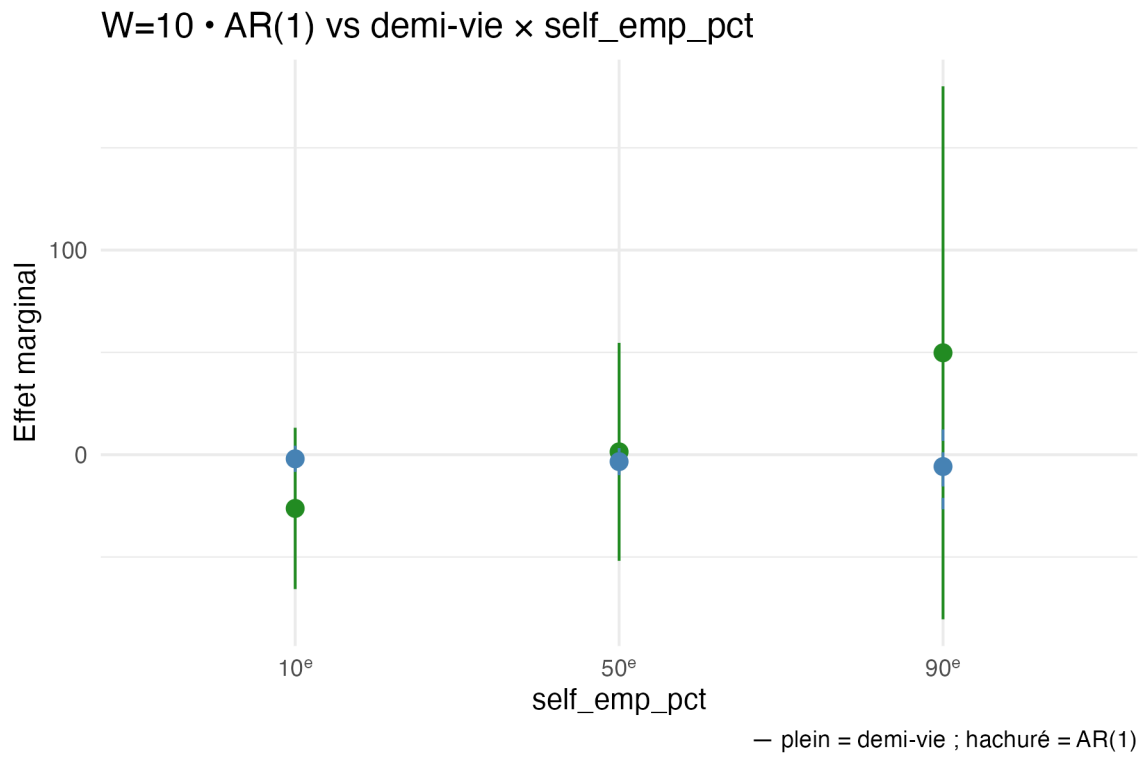
Graphique B.12 - W=10 effet marginal continu x self_emp_pct



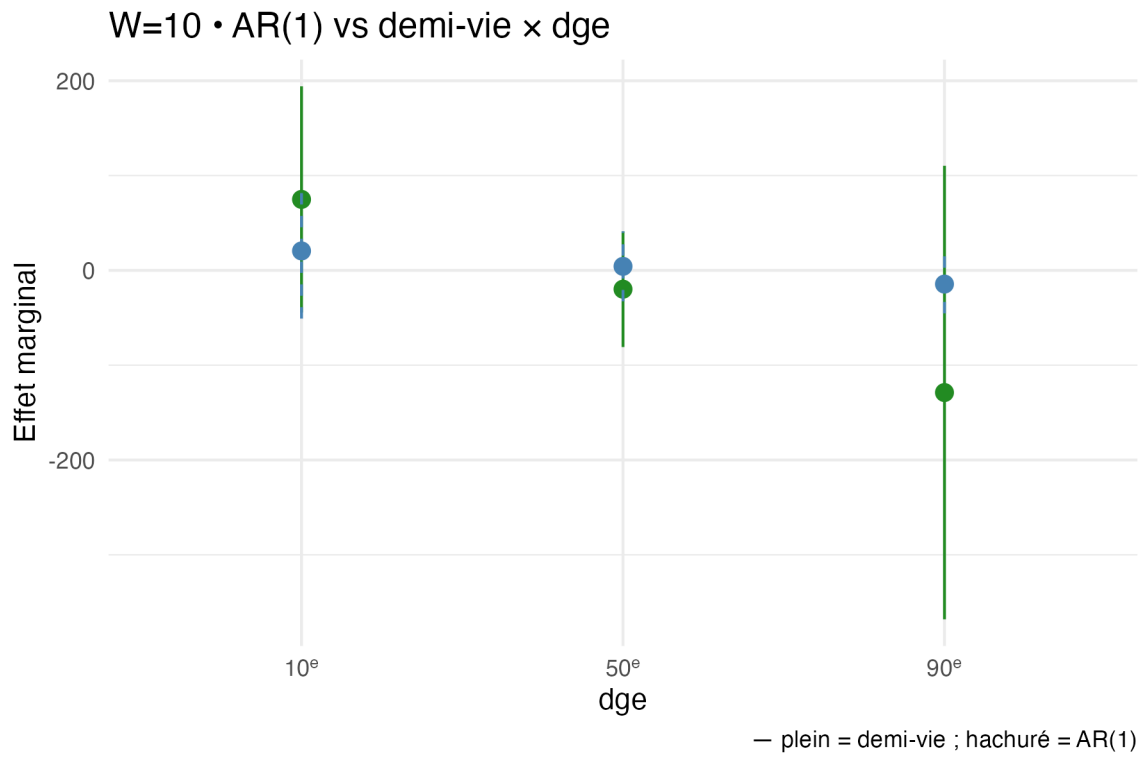
Graphique B.13 - W=5 effet marginal continu x self_emp_pct



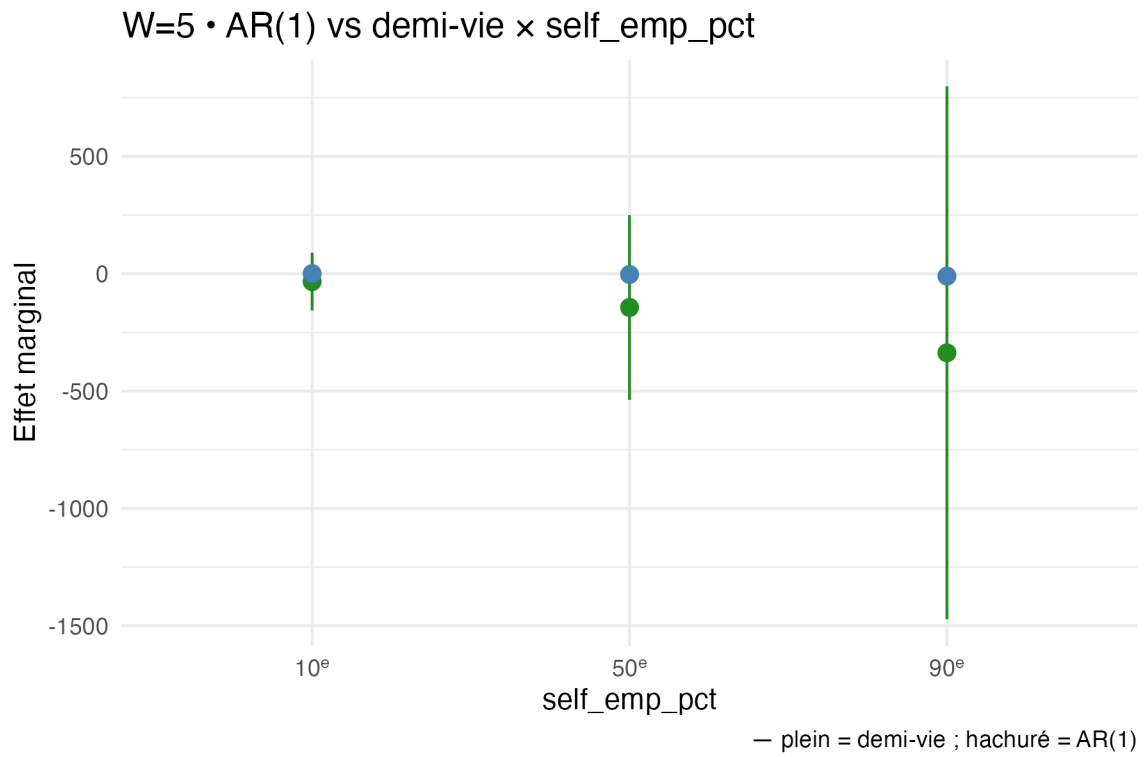
Graphique B.14 - W=10 AR(1) vs demi-vie x self_emp_pct



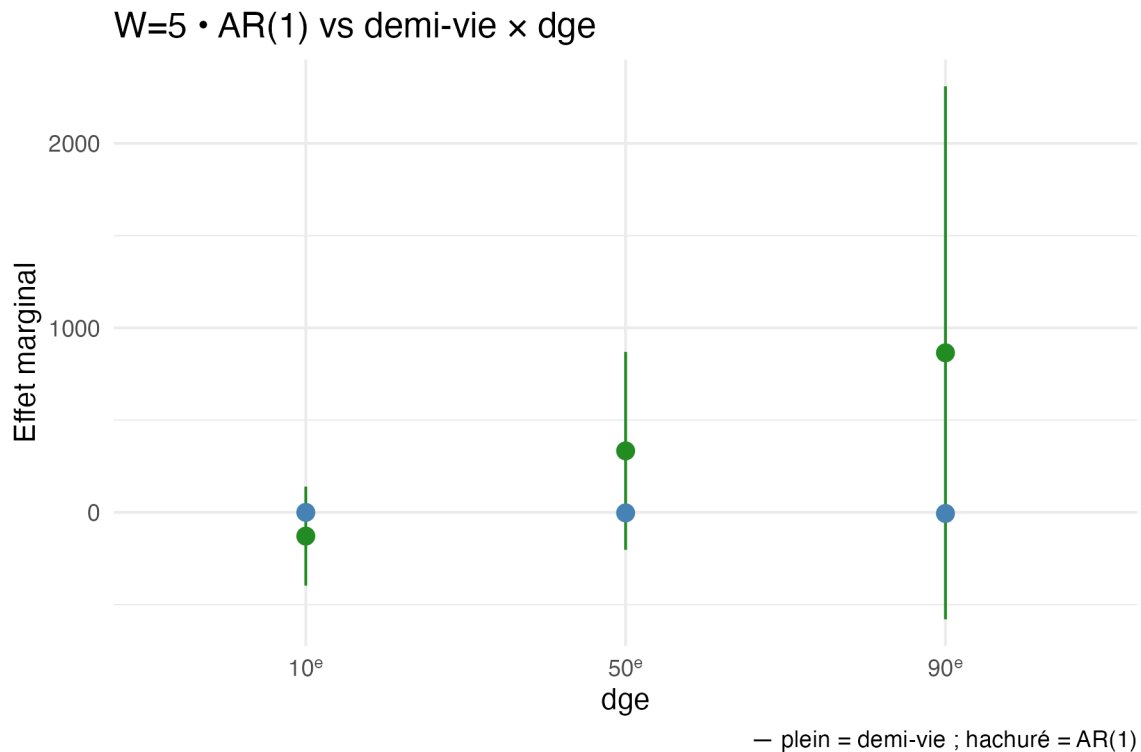
Graphique B.15 - W=10 AR(1) vs demi-vie x dge



Graphique B.16 - W=5 AR(1) vs demi-vie x self_emp_pct



Graphique B.17 - W=5 AR(1) vs demi-vie x dge



On propose d'autres visualisations : 1) Nuage de points AR(1)_W vs half_life_W coloré par le modérateur :

```
# 1) Scatter AR(1) vs demi-vie, coloré par modérateur
library(ggplot2)
library(dplyr)
for (W in c(5,10)) {
  df <- if (W==5) df5 else df10
  ar1 <- paste0("ar1_", W)
  hl <- paste0("half_life", W)
  for (mod in c("dge","self_emp_pct")) {
    p <- df %>%
      filter(!is.na(.data[[ar1]]), !is.na(.data[[hl]]), !is.na(.data[[mod]]))
    p <- p %>%
      ggplot(aes(
        x = .data[[ar1]],
        y = .data[[hl]],
        color = .data[[mod]]
      )) +
      geom_point(alpha = 0.6) +
      scale_color_viridis_c(name = mod) +
```

```

    labs(
      title = paste0("W=", W, " • AR(1) vs demi-vie coloré par ", mod),
      x = paste0("AR(1) (W=", W, ")"),
      y = paste0("Demi-vie (W=", W, ")")
    ) +
    theme_minimal()
  print(p)
}
}

library(ggplot2)
library(dplyr)

# 2) Téléchargement
out_dir <- "/Users/.../Chap3/scatter_AR1_vs_half_life"
if (!dir.exists(out_dir)) dir.create(out_dir, recursive = TRUE)

for (W in c(5, 10)) {
  df <- if (W == 5) df5 else df10
  ar1 <- paste0("ar1_", W)
  hl <- paste0("half_life", W)

  for (mod in c("dge", "self_emp_pct")) {
    # Filtrer les observations complètes
    df_plot <- df %>%
      filter(!is.na(.data[[ar1]]), !is.na(.data[[hl]]), !is.na(.data[[mod]]))

    # Construire le plot
    p <- ggplot(df_plot, aes(
      x = .data[[ar1]],
      y = .data[[hl]],
      color = .data[[mod]]
    )) +
      geom_point(alpha = 0.6) +
      scale_color_viridis_c(name = mod) +
      labs(
        title = paste0("W=", W, " • AR(1) vs demi-vie coloré par ", mod),
        x = paste0("AR(1) (W=", W, ")"),
        y = paste0("Demi-vie (W=", W, ")")
      ) +
      theme_minimal()

    # Affichage à l'écran

```

```

print(p)

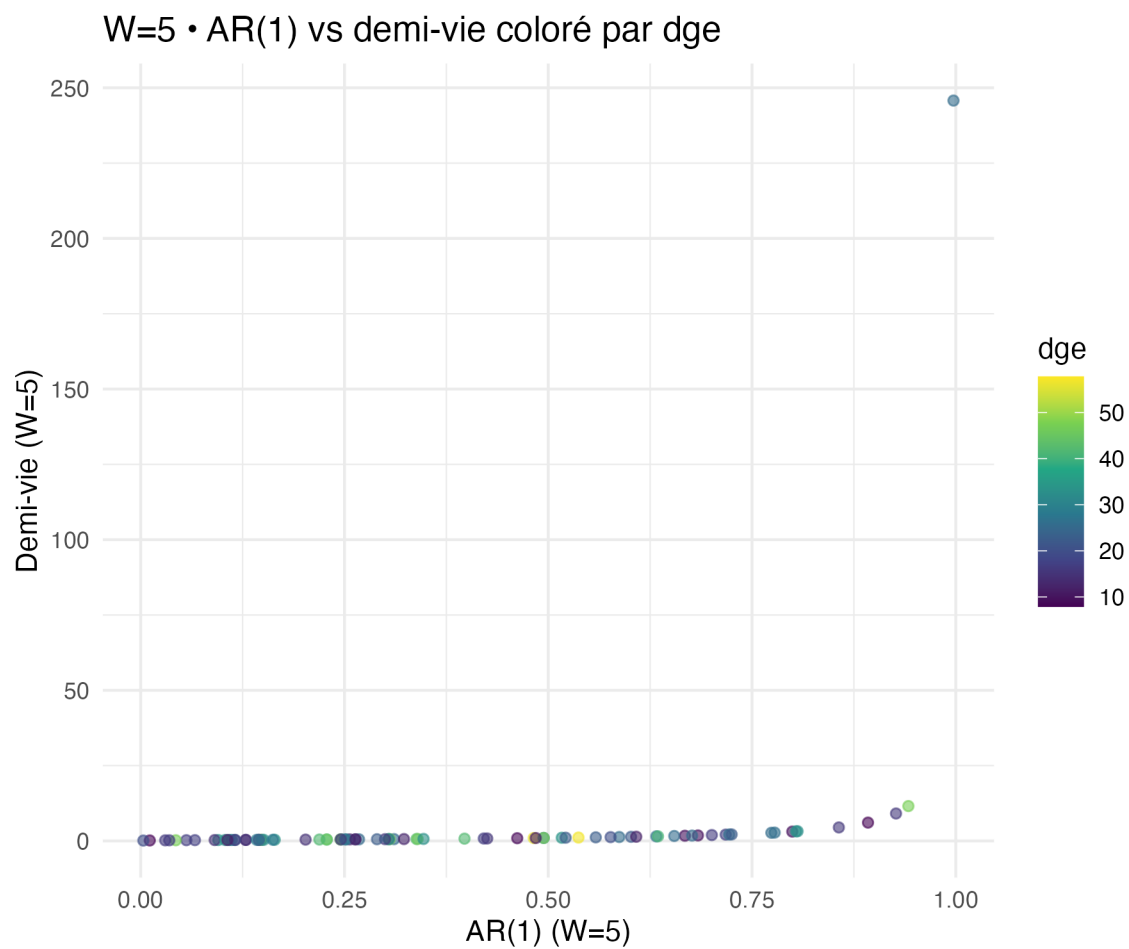
# Enregistrement
fname <- file.path(
  out_dir,
  paste0("W", W, "_scatter_AR1_vs_half_life_by_", mod, ".png")
)
ggsave(
  filename = fname,
  plot     = p,
  width    = 6,
  height   = 5,
  dpi      = 300,
  bg       = "white"
)
}
}

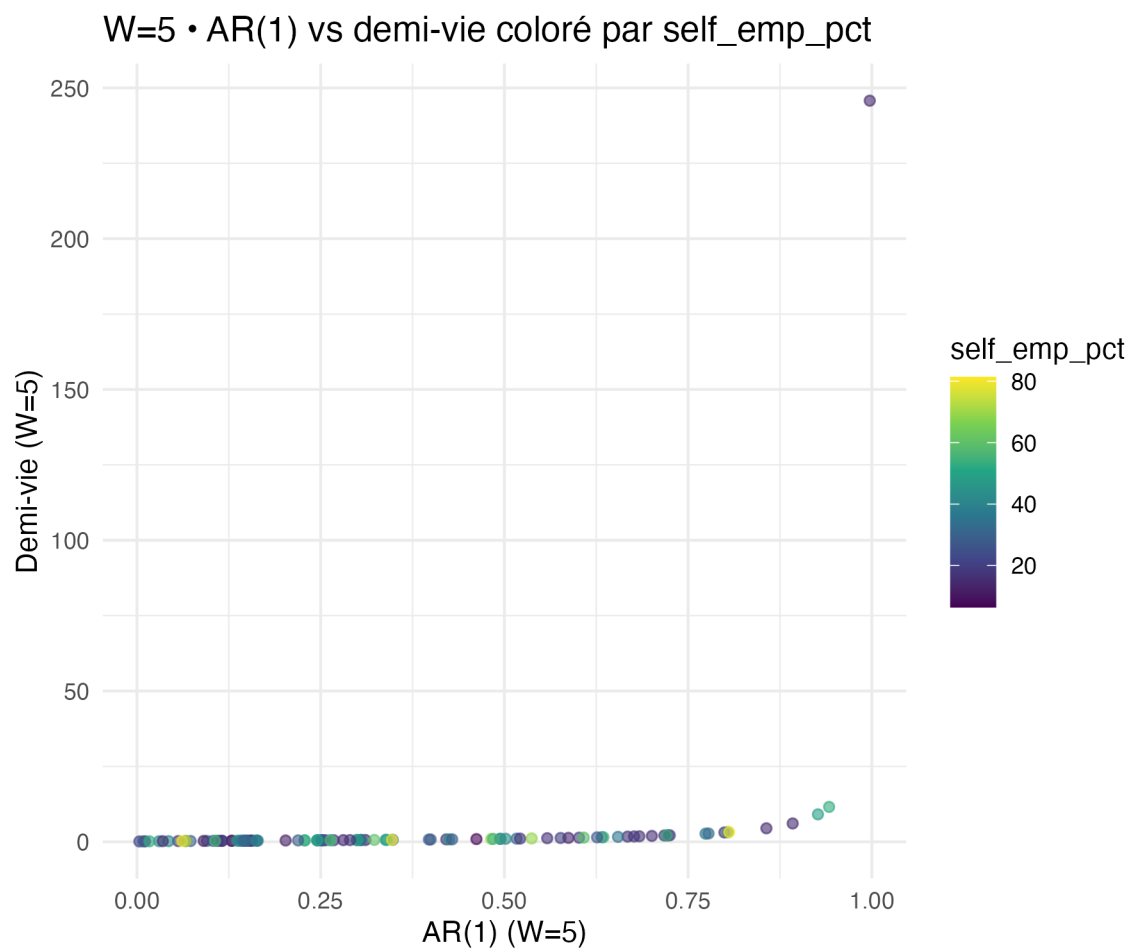
message("Tous les scatterplots ont été enregistrés dans : ", out_dir)

```

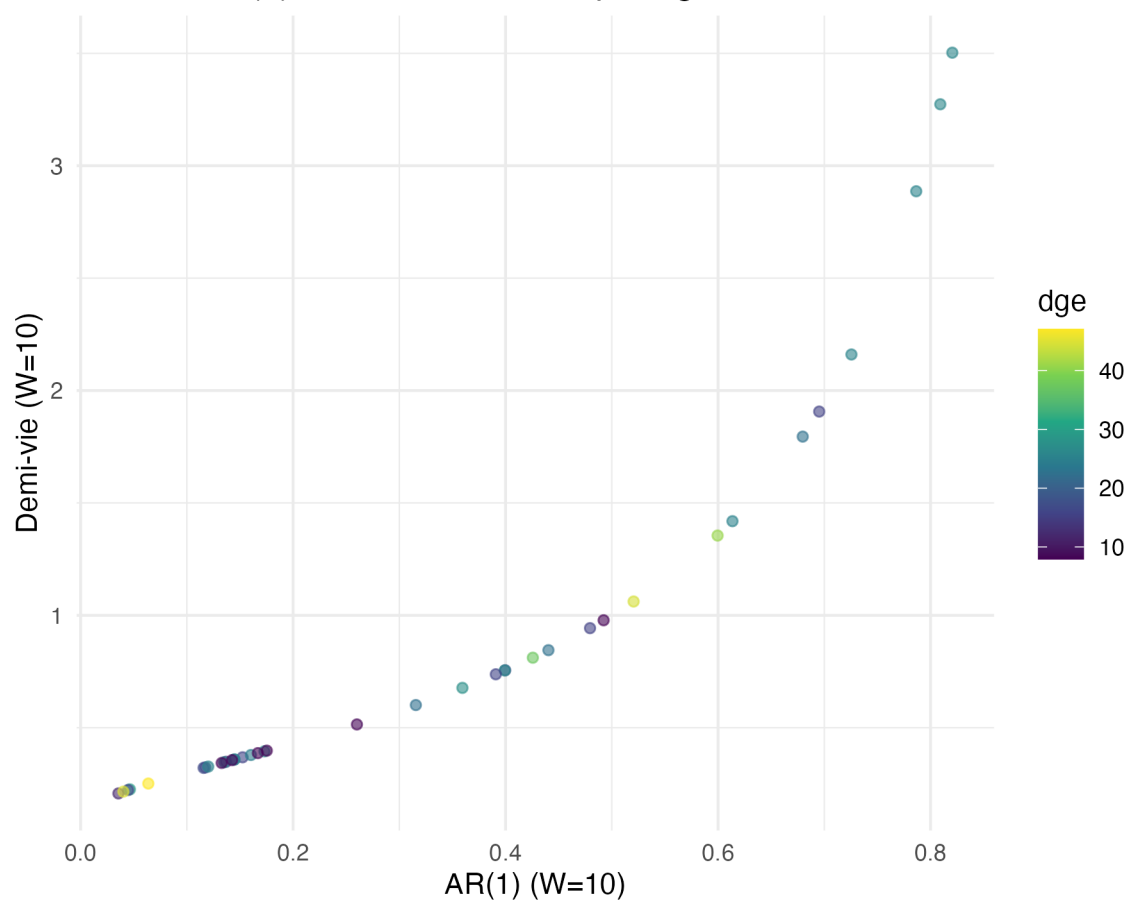
Les scatterplots comparant les coefficients AR(1) et les demi-vies de la croissance révèlent une **forte corrélation positive** entre ces deux mesures, quel que soit l'horizon temporel retenu (5 ou 10 ans). Plus le processus autorégressif est marqué, plus le temps nécessaire pour résorber un choc de croissance est long, ce qui confirme que ces deux indicateurs captent une dynamique commune de persistance macroéconomique. Cependant, l'introduction de la couleur pour représenter l'informalité, qu'elle soit mesurée par l'indice DGE ou par la part d'emploi informel (`self_emp_pct`), ne permet pas d'identifier un **motif clair ou structurant**. Dans le cas de DGE, les gradients de couleur apparaissent largement dispersés sur les nuages de points, sans regroupement systématique selon les niveaux d'informalité. Les pays fortement informels ne se situent ni préférentiellement en haut ni en bas de la diagonale, tant à court qu'à moyen terme. De même, la part d'emploi informel ne structure pas la relation entre AR(1) et demi-vie : si quelques cas extrêmes apparaissent pour les niveaux très élevés de `self_emp_pct`, ils ne dessinent pas une tendance cohérente. Dans l'ensemble, les couleurs se répartissent de façon hétérogène, sans former de couloir ou de trajectoire visible. Ces résultats suggèrent que ni DGE ni `self_emp_pct` **n'expliquent directement la relation conjointe** entre les deux mesures de persistance. Cela corrobore l'idée que l'effet modérateur de l'informalité, observé précédemment via les effets marginaux de GreenPotSharew, ne s'exprime pas sous la forme d'une translation simple ou d'une intensification linéaire de la persistance globale. Il agit plutôt en modulant **l'effet marginal conditionnel** du verdissement de l'emploi sur la dynamique de croissance, effet que l'on ne peut visualiser directement dans un scatterplot non paramétrique. L'informalité, en ce sens, ne transforme pas la structure intrinsèque de la persistance, mais conditionne la manière dont la transition verte agit sur elle.

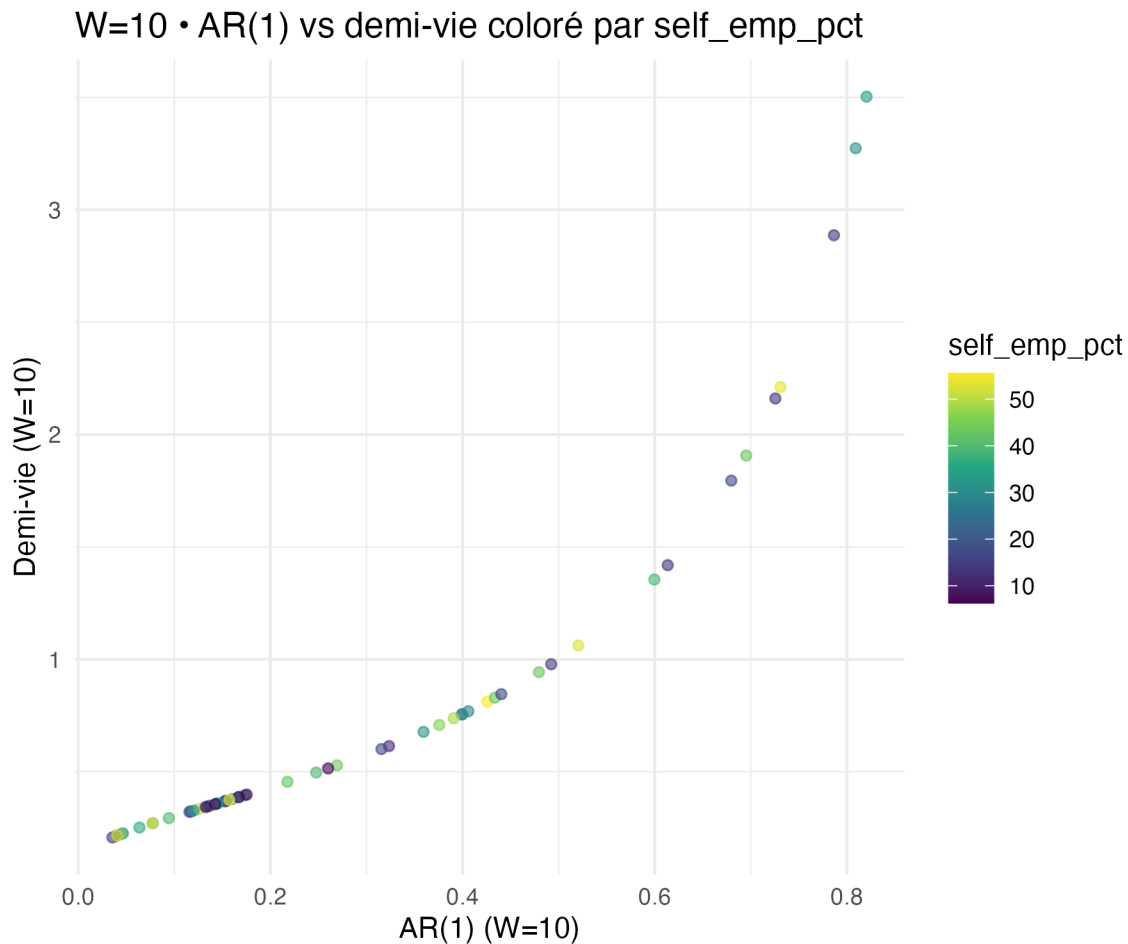
Graphique B.18 - Suite graphiques W=10 et W=5 AR(1) vs demi-vie colorés par dge ou self_emp_pct





W=10 • AR(1) vs demi-vie coloré par dge





2) Distribution (boxplots) de AR(1)_W et half_life_W par quartile du modérateur :

```
# 1) Boxplots par quartile du modérateur
library(ggplot2)
library(dplyr)
for (W in c(5,10)) {
  df <- if (W==5) df5 else df10
  ar1 <- paste0("ar1_", W)
  hl <- paste0("half_life", W)
  for (mod in c("dge","self_emp_pct")) {
    df_q <- df %>%
      filter(!is.na(.data[[mod]])) %>%
      mutate(
        quart = ntile(.data[[mod]], 4) %>% factor(labels = paste0("Q",1:4))
      )
  }
}
```

```

p1 <- ggplot(df_q, aes(x = quart, y = .data[[hl]])) +
  geom_boxplot(fill = "forestgreen", alpha = 0.5) +
  labs(
    title = paste0("W=",W," • Distrib. demi-vie par quartile de ",mod),
    x = paste0(mod," (quartiles)"),
    y = paste0("half_life",W)
  ) +
  theme_minimal()
p2 <- ggplot(df_q, aes(x = quart, y = .data[[ar1]])) +
  geom_boxplot(fill = "steelblue", alpha = 0.5) +
  labs(
    title = paste0("W=",W," • Distrib. AR(1) par quartile de ",mod),
    x = paste0(mod," (quartiles)"),
    y = paste0("ar1_",W)
  ) +
  theme_minimal()
print(p1); print(p2)
}
}

library(ggplot2)
library(dplyr)

# 2) Téléchargement
out_dir <- "/Users/.../Chap3/boxplots"
if (!dir.exists(out_dir)) dir.create(out_dir, recursive = TRUE)

for (W in c(5, 10)) {
  df <- if (W == 5) df5 else df10
  ar1 <- paste0("ar1_", W)
  hl <- paste0("half_life", W)

  for (mod in c("dge", "self_emp_pct")) {

    # découpage en quartiles
    df_q <- df %>%
      filter(!is.na(.data[[mod]])) %>%
      mutate(
        quart = ntile(.data[[mod]], 4) %>%
          factor(labels = paste0("Q", 1:4))
      )

    # boxplot demi-vie

```

```

p_hl <- ggplot(df_q, aes(x = quart, y = .data[[hl]])) +
  geom_boxplot(fill = "forestgreen", alpha = 0.5) +
  labs(
    title = paste0("W=", W, " • Distrib. demi-vie par quartile de ",
      ↪ mod),
    x      = paste0(mod, " (quartiles)"),
    y      = hl
  ) +
  theme_minimal()

# boxplot AR(1)
p_ar1 <- ggplot(df_q, aes(x = quart, y = .data[[ar1]])) +
  geom_boxplot(fill = "steelblue", alpha = 0.5) +
  labs(
    title = paste0("W=", W, " • Distrib. AR(1) par quartile de ", mod),
    x      = paste0(mod, " (quartiles)"),
    y      = ar1
  ) +
  theme_minimal()

# affichage à l'écran (optionnel)
print(p_hl)
print(p_ar1)

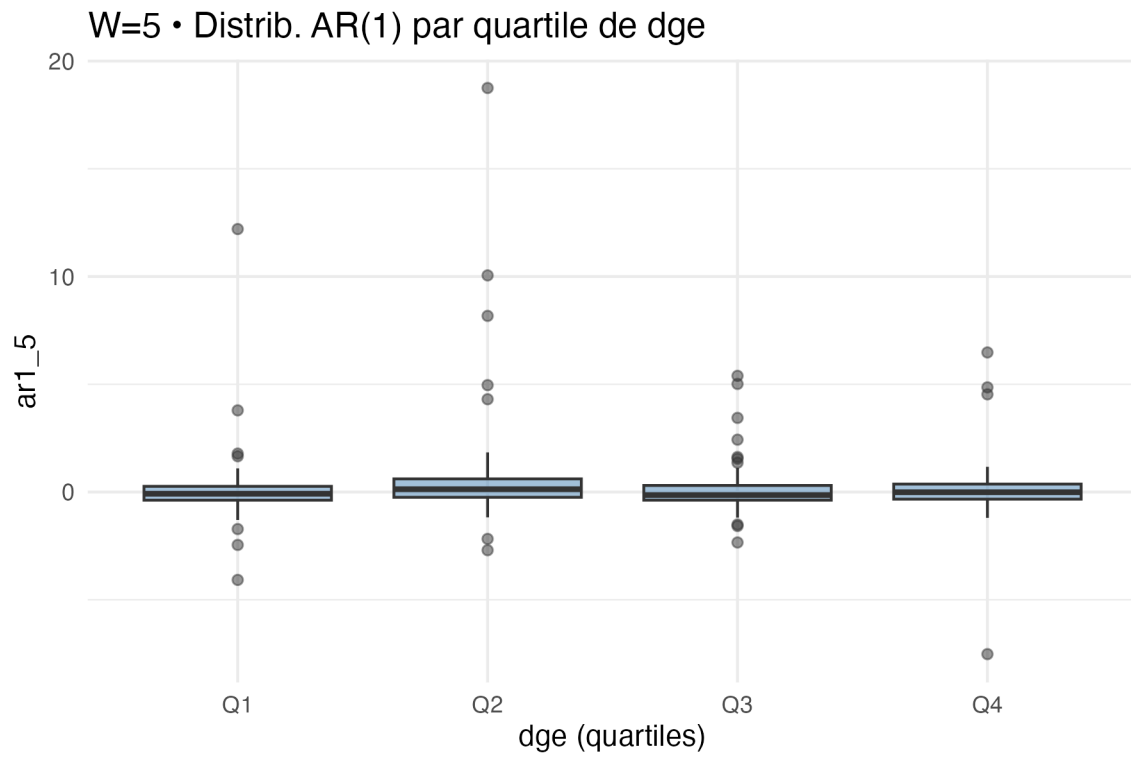
# enregistrement des fichiers
ggsave(
  filename = file.path(out_dir, paste0("W", W, "_half_life_by_", mod,
    ↪ "_boxplot.png")),
  plot      = p_hl,
  width     = 6, height = 4, dpi = 300, bg = "white"
)
ggsave(
  filename = file.path(out_dir, paste0("W", W, "_AR1_by_", mod,
    ↪ "_boxplot.png")),
  plot      = p_ar1,
  width     = 6, height = 4, dpi = 300, bg = "white"
)
}
}

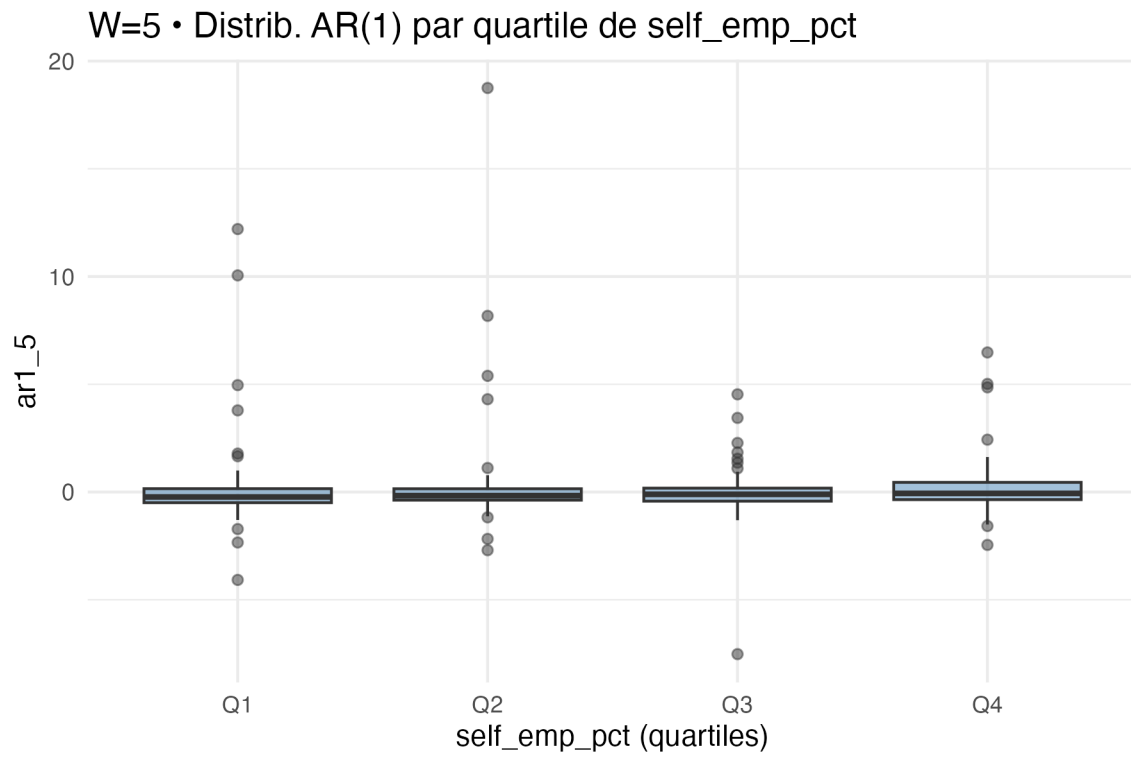
message("Tous les boxplots ont été enregistrés dans : ", out_dir)

```

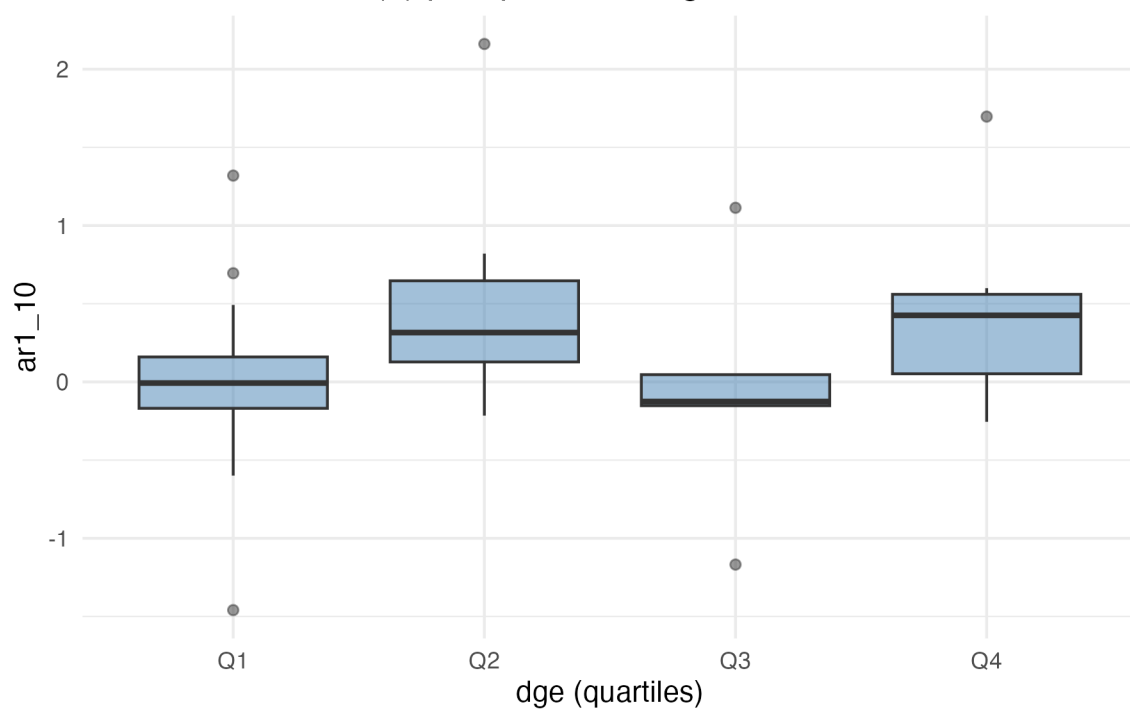
Les boxplots des indicateurs de persistance de la croissance (AR(1) et demi-vie), répartis par quartile d’informalité, révèlent que les effets de l’informalité sur la dynamique de croissance sont nettement plus visibles à moyen terme (fenêtre de 10 ans) qu’à court terme (5 ans). À horizon 10 ans, l’informalité institutionnelle (mesurée par DGE) semble structurer de manière non linéaire la persistance : le coefficient AR(1) présente une courbe en U, avec une baisse marquée en Q3 suivie d’un rebond en Q4, tandis que la demi-vie atteint des niveaux très élevés dans les Q2 et Q4, mais chute brutalement en Q3, suggérant des régimes de croissance très différenciés selon les configurations institutionnelles. Parallèlement, la part d’emploi informel (`self_emp_pct`) montre une progression plus régulière : les pays aux niveaux modérés ou élevés d’emploi informel (Q3 et Q4) présentent une persistance accrue, à la fois en AR(1) et en demi-vie, cette dernière culminant chez les pays les plus informels. Cela confirme que, dans certains contextes, l’informalité peut soutenir des cycles de croissance prolongés. À l’inverse, sur l’horizon court de 5 ans, les deux mesures restent largement centrées autour de zéro, quels que soient les niveaux d’informalité. Les quelques valeurs extrêmes observées (parfois très longues demi-vies ou AR(1) anormalement élevés) sont isolées et ne dessinent aucun motif robuste. Cela indique que les variations d’informalité ne produisent pas, à court terme, de différenciation claire en matière de persistance. Cette faible informativité de la fenêtre quinquennale confirme que les dynamiques de fond de la croissance, en lien avec l’informalité, se jouent principalement à l’échelle du moyen terme. En définitive, ces résultats confortent ceux des régressions interactives : l’effet structurant de l’informalité sur la persistance de la croissance, et donc sur la capacité des économies à absorber ou prolonger les effets de la transition verte, ne s’observe qu’à partir d’un certain horizon temporel. Les pays très informels peuvent alors se situer à l’un ou l’autre extrême, selon que l’informalité agit comme un amortisseur ou un facteur de fragilité. L’horizon de 10 ans apparaît ainsi comme le niveau d’analyse pertinent pour évaluer le rôle différencié de l’informalité dans les trajectoires de croissance durable.

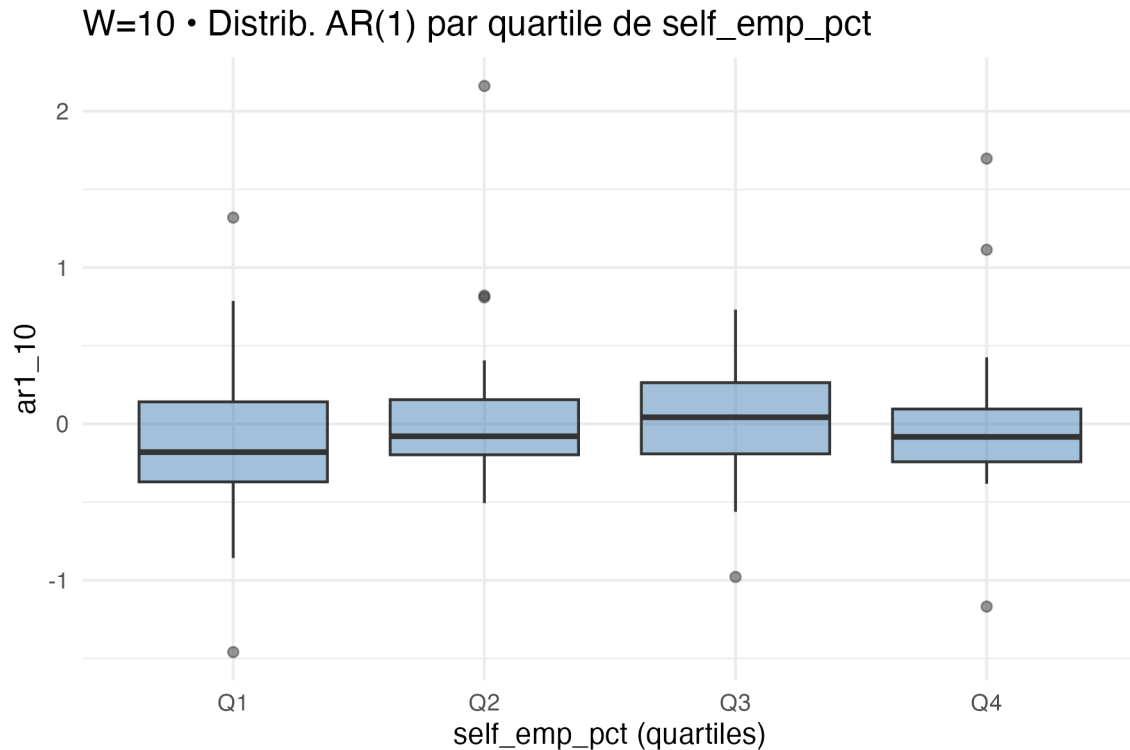
Graphique B.19 - Suite de boxplots de distribution demi-vie ou AR(1) par dge ou `self_emp_pct`





W=10 • Distrib. AR(1) par quartile de dge





3) Séries temporelles moyennées de AR(1)_W et half_life_W par quartile du modérateur :

```
# 1) Évolution temporelle de AR(1) et demi-vie par quartile
library(ggplot2)
library(dplyr)
library(tidyr)

for (W in c(5,10)) {
  df <- if (W==5) df5 else df10
  ar1 <- paste0("ar1_", W)
  hl <- paste0("half_life", W)

  for (mod in c("dge", "self_emp_pct")) {
    df_ts <- df %>%
      filter(!is.na(.data[[ar1]]),
             !is.na(.data[[hl]]),
             !is.na(.data[[mod]])) %>%
    mutate(
      q = ntile(.data[[mod]], 4) %>%

```

```

        factor(labels = paste0("Q",1:4))
    ) %>%
group_by(year, q) %>%
summarise(
    mean_ar1 = mean(.data[[ar1]], na.rm = TRUE),
    mean_hl  = mean(.data[[hl]],  na.rm = TRUE),
    .groups = "drop"
) %>%
pivot_longer(
    cols      = c(mean_ar1, mean_hl),
    names_to  = "indicator",
    values_to = "moyenne"
) %>%
mutate(
    indicator = case_when(
        indicator == "mean_ar1" ~ paste0("AR(1) (W=", W, ")"),
        indicator == "mean_hl"  ~ paste0("half-life (W=", W, ")")
    )
)

p <- ggplot(df_ts, aes(x = year, y = moyenne, color = q)) +
  geom_line(size = 1) +
  facet_wrap(~indicator, scales = "free_y") +
  labs(
    title = paste0("W=", W, " • Évolution temporelle par quartile de ",
      ↪ mod),
    x      = "Année",
    y      = "Moyenne",
    color  = paste0(mod, " (quartile)")
  ) +
  theme_minimal()

print(p)
}
}

library(ggplot2)
library(dplyr)
library(tidyr)

# 2) Téléchargement
out_dir <- "/Users/.../Chap3/time_series_by_quartile"
if (!dir.exists(out_dir)) dir.create(out_dir, recursive = TRUE)

```

```

for (W in c(5, 10)) {
  df <- if (W == 5) df5 else df10
  ar1 <- paste0("ar1_", W)
  hl <- paste0("half_life", W)

  for (mod in c("dge", "self_emp_pct")) {
    df_ts <- df %>%
      filter(!is.na(.data[[ar1]]),
             !is.na(.data[[hl]]),
             !is.na(.data[[mod]])) %>%
      mutate(
        q = ntile(.data[[mod]], 4) %>%
          factor(labels = paste0("Q", 1:4))
      ) %>%
      group_by(year, q) %>%
      summarise(
        mean_ar1 = mean(.data[[ar1]], na.rm = TRUE),
        mean_hl = mean(.data[[hl]], na.rm = TRUE),
        .groups = "drop"
      ) %>%
      pivot_longer(
        cols = c(mean_ar1, mean_hl),
        names_to = "indicator",
        values_to = "moyenne"
      ) %>%
      mutate(
        indicator = case_when(
          indicator == "mean_ar1" ~ paste0("AR(1) (W=", W, ")"),
          indicator == "mean_hl" ~ paste0("Demi-vie (W=", W, ")")
        )
      )

    p <- ggplot(df_ts, aes(x = year, y = moyenne, color = q)) +
      geom_line(size = 1) +
      facet_wrap(~indicator, scales = "free_y") +
      labs(
        title = paste0("W=", W, " • Évolution temporelle par quartile de ",
          ↪ mod),
        x = "Année",
        y = "Moyenne",
        color = paste0(mod, " (quartile)")
      ) +

```

```

    theme_minimal()

# Affichage
print(p)

# Enregistrement du graphique
fname <- file.path(
  out_dir,
  paste0("W", W, "_time_series_", mod, ".png")
)
ggsave(
  filename = fname,
  plot      = p,
  width     = 8,
  height    = 5,
  dpi       = 300,
  bg        = "white"
)
}
}

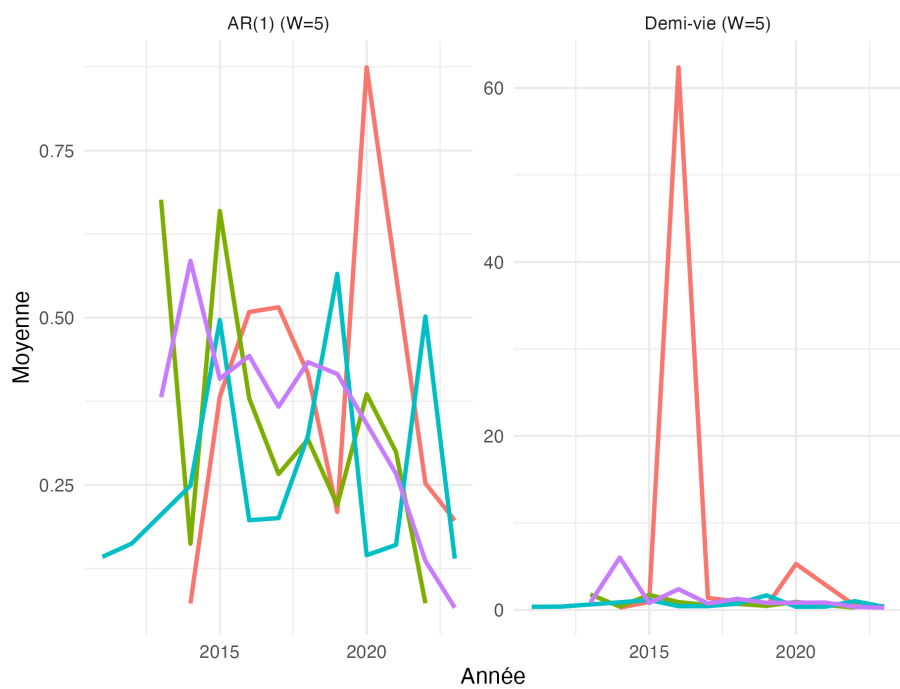
message("Tous les graphiques temporels ont été enregistrés dans : ", out_dir)

```

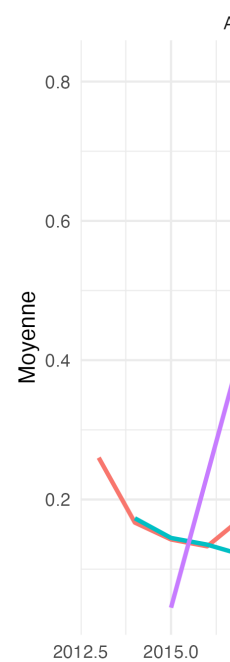
Les séries temporelles de la demi-vie et du coefficient AR(1) révèlent que la persistance de la croissance réagit de manière différenciée selon le niveau d’informalité et l’indicateur mobilisé. À court et moyen termes, des **pics de demi-vie très marqués mais isolés** apparaissent systématiquement dans le **troisième quartile** des modérateurs (`dge` et `self_emp_pct`), notamment en 2014–2016 et en 2018–2020. Ces sursauts traduisent une **sensibilité accrue** des pays à informalité intermédiaire face aux chocs conjoncturels. En revanche, l’**AR(1) évolue de façon plus régulière**, sans valeurs extrêmes, mais confirme lui aussi que le Q3 présente la plus forte persistance, particulièrement à 10 ans. Une **hiérarchie stable** se dégage : Q3 est le plus réactif, Q4 suit avec retard ou modération, tandis que Q1 et Q2 restent globalement stables. À 5 ans, la persistance reste faible dans l’ensemble, quel que soit le modérateur. Ces résultats confirment que les **effets de la transition verte sur la stabilité de la croissance dépendent fortement du niveau d’informalité**, avec une instabilité plus marquée dans les contextes intermédiaires.

Graphique B.20 - Suite graphiques évolutions temporelles par quartile

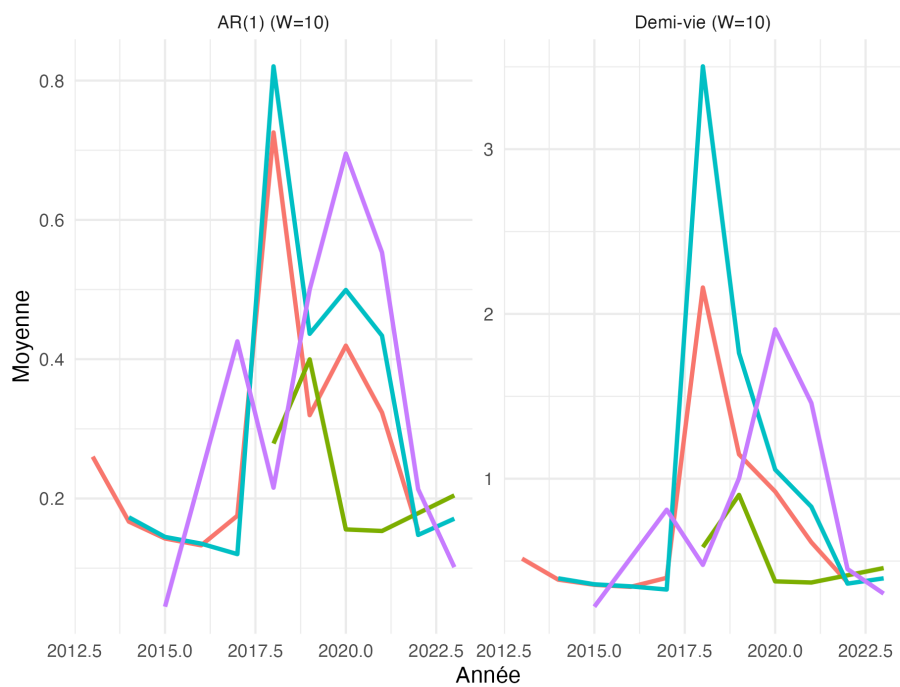
W=5 • Évolution temporelle par quartile de self_emp_pct



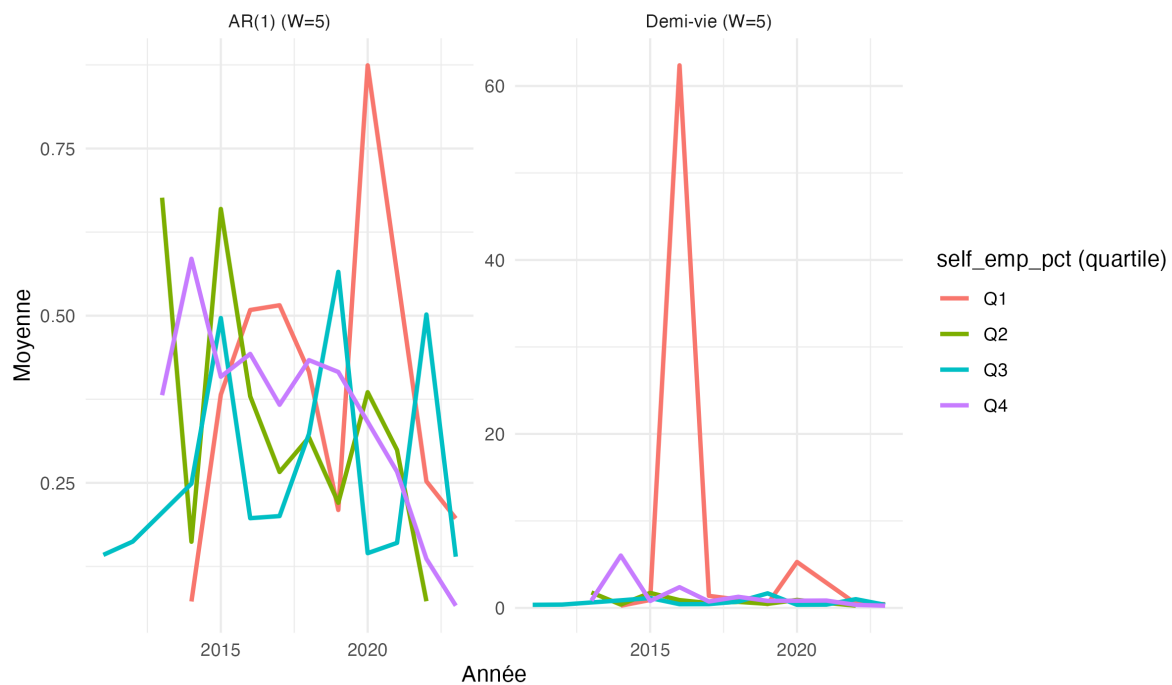
W=10 • Évol



W=10 • Évolution temporelle par quartile de self_emp_pct



W=5 • Évolution temporelle par quartile de self_emp_pct



6 Robustesses et extensions

6.1 Panel GMM dynamique (Arellano–Bond one-step et two-step, puis System-GMM)

```
library(plm)
library(pgmm)
library(dplyr)

# 1) Fonction utilitaire pour ajouter des étoiles de significativité
add_stars <- function(pvals) {
  symnum(pvals,
    cutpoints = c(0, 0.001, 0.01, 0.05, 0.1, 1),
    symbols   = c(" ", " ", " ", " .", " "))
}

# 2) Préparer l'objet panel
pdata <- pdata.frame(df1_mod1, index = c("iso3c", "year"))
```

```

# 3) One-step Arellano-Bond différencié avec collapse
gmm1 <- pgmm(
  DeltaY ~ lag(DeltaY, 1)
    + GreenPotSharew * dge
    + deltaLnK + DeltaLnL + ays
    + gov_exp_pctgdp + open_pct + access_electricity
  | lag(DeltaY, 2:5),
  data      = pdata,
  effect     = "individual",
  model      = "onestep",
  transformation = "d",
  collapse   = TRUE
)

# extraction de estimates + SE
b1 <- coef(gmm1)
v1 <- vcov(gmm1)
se1 <- sqrt(diag(v1))
p1 <- 2 * pnorm(-abs(b1 / se1))

res1 <- tibble(
  Variable = names(b1),
  Estimate = b1,
  Std.Error = se1,
  t.value = b1 / se1,
  p.value = p1
) %>%
  mutate(Signif = add_stars(p.value))

print(res1)

# tests de spécification
cat("\nAR(1) in residuals:\n"); mtest(gmm1, order = 1)
cat("\nAR(2) in residuals (should be NS):\n"); mtest(gmm1, order = 2)
cat("\nHansen overid test:\n"); mtest(gmm1, type = "Hansen")

# 4) Two-step Arellano-Bond
gmm2 <- update(gmm1, model = "twosteps")
b2 <- coef(gmm2)
v2 <- vcov(gmm2)
se2 <- sqrt(diag(v2))

```



```

p2    <- 2 * pnorm(-abs(b2 / se2))

res2 <- tibble(
  Variable   = names(b2),
  Estimate   = b2,
  Std.Error  = se2,
  t.value    = b2 / se2,
  p.value    = p2
) %>%
  mutate(Signif = add_stars(p.value))

print(res2)
cat("\nAR(2) residuals (2-step):\n");    mtest(gmm2, order = 2)
cat("\nHansen overid (2-step):\n");      mtest(gmm2, type = "Hansen")

# 5) System-GMM two-step (niveau + diff)
sysgmm <- pgmm(
  DeltaY ~ lag(DeltaY, 1)
    + GreenPotSharew * dge
    + deltaLnK + DeltaLnL + ays
    + gov_exp_pctgdp + open_pct + access_electricity
| lag(DeltaY, 2:5)    # instruments pour diffs
  + lag(DeltaY, 0:1), # instruments pour niveaux
data      = pdata,
effect    = "individual",
model     = "twosteps",
transformation = "ld",
collapse  = TRUE
)

bsys    <- coef(sysgmm)
vsys    <- vcov(sysgmm)
sesys   <- sqrt(diag(vsys))
psys    <- 2 * pnorm(-abs(bsys / sesys))

resys <- tibble(
  Variable   = names(bsys),
  Estimate   = bsys,
  Std.Error  = sesys,
  t.value    = bsys / sesys,
  p.value    = psys
) %>%

```

```
mutate(Signif = add_stars(p.value))

print(resys)
cat("\nAR(2) residuals (System):\n"); mtest(sysgmm, order = 2)
cat("\nHansen overid (System):\n"); mtest(sysgmm, type = "Hansen")
```

Les résultats des modèles GMM montrent une forte hétérogénéité selon la spécification. Le modèle Arellano-Bond en différences (1-step et 2-step) ne détecte aucun effet significatif ni de la part d'emplois verts (GreenPotSharew), ni de son interaction avec l'informalité (dge). En revanche, dans le modèle **System-GMM**, l'effet direct de GreenPotSharew devient **positif et hautement significatif**, tout comme l'interaction avec dge, dont le coefficient est **négatif et significatif**. Cela indique que l'effet bénéfique des EFPV sur la croissance est atténué dans les contextes très informels. Ce même modèle montre également la significativité attendue des contrôles standards (capital, travail, ouverture commerciale, dépenses publiques), renforçant la crédibilité des estimations. Les tests de spécification confirment l'absence d'auto-corrélation de second ordre (AR(2)) dans tous les modèles, mais soulignent une **légère sur-identification** dans le 2-step Arellano-Bond. Au total, seul le modèle System-GMM produit des résultats cohérents et statistiquement robustes, confirmant que **la transition verte stimule la croissance**, mais que **cet effet décroît avec le niveau d'informalité**.

Tableau B.11 - Arellano-Bond One-Step GMM (Differenced, collapse = TRUE)

| Variable | Estimate | Std. Error | t value | p-value |
|-----------------------------|------------|------------|---------|---------|
| lag(DeltaY, 1) | 0.0266223 | 0.2797904 | 0.0952 | 0.9242 |
| GreenPotSharew | 0.0305334 | 0.7054943 | 0.0433 | 0.9655 |
| dge | 0.0119299 | 0.0278167 | 0.4289 | 0.6680 |
| deltalnK | 0.1080474 | 0.0724289 | 1.4918 | 0.1358 |
| DeltaLnL | 0.3071335 | 0.4327574 | 0.7097 | 0.4779 |
| ays | -0.0162842 | 0.0410086 | -0.3971 | 0.6913 |
| gov_exp_pctgdp | -0.0048721 | 0.0037230 | -1.3087 | 0.1907 |
| open_pct | 0.0017545 | 0.0016886 | 1.0390 | 0.2988 |
| access_electricity | 0.0002732 | 0.0039399 | 0.0693 | 0.9447 |
| GreenPotSharew × dge | 0.0014961 | 0.0371103 | 0.0403 | 0.9678 |

Tableau B.12 - Arellano-Bond Two-Step GMM

| Variable | Estimate | Std. Error | t value | p-value |
|----------------|-----------|------------|---------|---------|
| lag(DeltaY, 1) | 0.0137951 | 0.0755966 | 0.1825 | 0.8552 |
| GreenPotSharew | 0.0347942 | 0.0468898 | 0.7420 | 0.4581 |
| dge | 0.0122305 | 0.0057065 | 2.1432 | 0.0321* |

| Variable | Estimate | Std. Error | t value | p-value |
|-----------------------------|------------|------------|---------|-------------|
| deltalnK | 0.1124481 | 0.0189919 | 5.9209 | < 0.0001*** |
| DeltaLnL | 0.3583751 | 0.0868960 | 4.1242 | < 0.0001*** |
| ays | -0.0128009 | 0.0057754 | -2.2165 | 0.0267* |
| gov_exp_pctgdp | -0.0046514 | 0.0009352 | -4.9738 | < 0.0001*** |
| open_pct | 0.0020254 | 0.0004445 | 4.5570 | < 0.0001*** |
| access_electricity | 0.0007559 | 0.0007571 | 0.9984 | 0.3181 |
| GreenPotSharew \times dge | 0.0009885 | 0.0024701 | 0.4002 | 0.6890 |

Tableau B.13 - System-GMM Two-Step (Levels + Differences)

| Variable | Estimate | Std. Error | t value | p-value |
|---|------------|------------|---------|-------------|
| lag(DeltaY, 1) | 0.0046428 | 0.0497068 | 0.0934 | 0.9256 |
| GreenPotSharew | 0.1037118 | 0.0130711 | 7.9344 | < 0.0001*** |
| dge | 0.0012154 | 0.0002335 | 5.2052 | < 0.0001*** |
| deltalnK | 0.1850217 | 0.0134819 | 13.7237 | < 0.0001*** |
| DeltaLnL | 0.6230491 | 0.0728244 | 8.5555 | < 0.0001*** |
| ays | 0.0006412 | 0.0008494 | 0.7549 | 0.4503 |
| gov_exp_pctgdp | -0.0008559 | 0.0001392 | -6.1479 | < 0.0001*** |
| open_pct | 0.0000965 | 0.0000248 | 3.8918 | 0.0001*** |
| · access_electricity | -0.0002020 | 0.0001151 | -1.7552 | 0.0792 · |
| GreenPotSharew \times dge | -0.0023169 | 0.0006748 | -3.4335 | 0.000596*** |

Tableau B.14 - Spécification et validité des instruments

| Test | 1-step AB | 2-step AB | System-GMM |
|---------------------------------------|--------------|-----------------------|--------------|
| AR(1) différencié (doit être présent) | $p = 0.2645$ | – | – |
| AR(2) différencié (doit être NS) | $p = 0.5569$ | $p = 0.4613$ | $p = 0.4614$ |
| Hansen overid. | $p = 0.2645$ | $p = 0.0343 \uparrow$ | $p = 0.0639$ |

- Absence d'AR(2) dans les résidus partout ($p > 0.1$) \rightarrow les instruments en différences sont valides.
- Hansen test :
 - 1-step : p 0.26 instruments globalement valides.
 - 2-step : p 0.03 légère sous-identification (on pourrait alléger le jeu d'instruments).
 - System-GMM : p 0.06 acceptable.

6.2 LSDV dynamique (PLM «Within» avec correction Arellano)

```
library(plm)      # pour plm()
library(lmtest)   # pour coeftest()
library(sandwich) # pour vcovHC()

# 1) Formule dynamique (tout sur une ligne)
dyn_formula <- DeltaY ~ lag(DeltaY, 1) +
  GreenPotSharew * dge +
  deltaInK + DeltaLnL + ays +
  gov_exp_pctgdp + open_pct + access_electricity

# 2) Estimation "within" (LSDV dynamique)
plm_dyn <- plm(
  formula = dyn_formula,
  data    = df1_mod1,
  index   = c("iso3c", "year"),
  model   = "within"
)

# 3) Résumé standard (sans correction)
summary(plm_dyn)

# 4) Matrice de covariance corrigée Arellano (HC + auto-corrélation)
arellano_se <- vcovHC(
  plm_dyn,
  method = "arellano", # correction Arellano (1987)
  type    = "HC1",      # hétéroscédasticité robuste
  cluster = "group"     # clustering par pays
)

# 5) Tests des coefficients avec SE corrigés
coeftest(plm_dyn, vcov = arellano_se)
```

Dans ce modèle dynamique à effets fixes (LSDV) corrigé des biais de Nickell via l'erreur standard HC1 et un clustering par pays, ni la part d'EFPV ni son interaction avec l'informalité (dge) n'ont d'effet significatif sur la dynamique de croissance. Le terme de persistance inertielle (ΔY) est légèrement négatif mais non significatif, indiquant l'absence d'effet d'entraînement de la croissance passée. L'effet direct de GreenPotSharew est négatif, et son interaction avec dge est faiblement positive, mais aucun des deux n'atteint un seuil de significativité. En revanche, les variables de contrôle macroéconomiques jouent un rôle important : l'accumulation de capital et la variation de l'emploi ont un impact positif et très significatif sur la croissance,

tandis que le niveau moyen de scolarité et les dépenses publiques en pourcentage du PIB ont un effet négatif robuste. L'ouverture commerciale montre un effet légèrement positif, tandis que l'accès à l'électricité ne joue pas significativement. En somme, dans ce cadre dynamique, ce sont les fondamentaux économiques classiques, capital, emploi, éducation, politique budgétaire, qui structurent la croissance, sans que l'effet des emplois verts ni leur interaction avec l'informalité ne se distingue de manière robuste.

Tableau B.15 - Estimation « within » standard (sans correction)

| Variable | Estimate | Std. Error | t value | Pr(> t) | Signif. |
|----------------------|------------|------------|---------|-----------------------|---------|
| lag(DeltaY, 1) | -0.0549967 | 0.040878 | -1.345 | 0.1798 | |
| GreenPotSharew | -0.0865238 | 0.122532 | -0.706 | 0.4808 | |
| dge | -0.0029429 | 0.003376 | -0.872 | 0.3842 | |
| deltalnK | 0.1344637 | 0.012547 | 10.717 | $< 2 \times 10^{-1}$ | *** |
| DeltaLnL | 0.3689951 | 0.075563 | 4.883 | 1.97×10^{-4} | *** |
| ays | -0.0183935 | 0.005014 | -3.669 | 3.04×10^{-4} | *** |
| gov_exp_pctgdp | -0.0058678 | 0.000603 | -9.739 | $< 2 \times 10^{-1}$ | *** |
| open_pct | 0.0003352 | 0.000184 | 1.822 | 0.0698 | |
| access_electricity | -0.0005189 | 0.000425 | -1.221 | 0.2233 | |
| GreenPotSharew × dge | 0.0052604 | 0.006415 | 0.820 | 0.4131 | |

Tableau B.16 - Test des coefficients avec SE corrigés (Arellano HC1, clustering par pays)

| Variable | Estimate | Std. Error | t value | Pr(> t) | Signif. |
|----------|----------|------------|---------|----------|---------|
|----------|----------|------------|---------|----------|---------|

| | Estimate | Lower CI | Upper CI | Wald Stat | p-value |
|-----------------------------|------------|----------|----------|------------------------|---------|
| lag(DeltaY, 1) | -0.0549967 | 0.063298 | -0.869 | 0.3858 | |
| GreenPotSharew | -0.0865238 | 0.122785 | -0.705 | 0.4817 | |
| dge | -0.0029429 | 0.006398 | -0.460 | 0.6460 | |
| deltalnK | 0.1344637 | 0.021019 | 6.397 | 8.93×10^{-1} | *** |
| DeltaLnL | 0.3689951 | 0.105366 | 3.502 | 0.000556 | *** |
| ays | -0.0183935 | 0.006435 | -2.858 | 0.004655 | ** |
| gov_exp_pctgdp | -0.0058678 | 0.000796 | -7.375 | 3.05×10^{-12} | *** |
| open_pct | 0.0003352 | 0.000205 | 1.634 | 0.1035 | |
| access_electricity | -0.0005189 | 0.000995 | -0.521 | 0.6027 | |
| GreenPotSharew × dge | 0.0052604 | 0.007571 | 0.695 | 0.4879 | |

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ · $p < 0.1$

6.3 Anderson–Hsiao IV-LSDV

```
library(dplyr)
library(fixest)

# 1) Spécification IV-LSDV
ah_model <- feols(
  # y ~ RHS regressors | FE | IV
  DeltaY ~
    GreenPotSharew * dge
  + deltalnK + DeltaLnL + ays
  + gov_exp_pctgdp + open_pct + access_electricity
```

```

| iso3c                                # 1) FE pays
| lag(DeltaY, 1) ~ lag(DeltaY, 2)      # 2) IV: endogène ~ instrument
, cluster = ~iso3c                     # 3) clustering
, data    = df1_mod1
)

# 2) Résumé du modèle IV-LSDV
summary(ah_model)

# 3) Première étape séparée (pour le test de force de l'instrument)
first_stage <- feols(
  lag(DeltaY, 1) ~ lag(DeltaY, 2)
  + deltalnK + DeltaLnL + ays
  + gov_exp_pctgdp + open_pct + access_electricity
| iso3c,
  data    = df1_mod1,
  cluster = ~iso3c
)

# 4) Extraction de F-stat de Kleibergen-Paap
fs <- fitstat(first_stage, "F")
print(fs)

```

Dans ce modèle IV-LSDV avec effet fixe pays et erreurs clusterisées, l'instrument $\text{lag}(\Delta Y)$ pour $\text{lag}(\Delta Y)$ est extrêmement fort ($F = 5 \times 10^{32}$), ce qui garantit la qualité de l'instrumentation. Le coefficient ajusté de ΔY est parfaitement significatif, ce qui confirme une forte inertie dynamique. L'effet direct de la part d'emplois verts (GreenPotSharew) est très faible mais significatif ($p = 0,046$), tandis que l'effet de l'informalité seule (dge) est également significatif ($p = 0,0076$). L'interaction $\text{GreenPotSharew} \times \text{dge}$ est négative et significative ($p = 0,048$), ce qui indique que plus l'informalité est élevée, plus l'impact des emplois verts sur la croissance diminue. Cela suggère que l'informalité agit comme un frein à l'effet stabilisateur potentiel de la transition verte. Par ailleurs, certaines variables de contrôle conservent une influence robuste : le capital (deltalnK) a un effet négatif significatif, les dépenses publiques (gov_exp_pctgdp) et l'accès à l'électricité ont un effet positif et significatif, tandis que l'emploi et l'ouverture commerciale n'ont pas d'effet statistiquement discernable. En résumé, ce modèle IV met en évidence un effet modérateur significatif de l'informalité sur la contribution des emplois verts à la croissance, confirmant que la structure institutionnelle du marché du travail conditionne la soutenabilité de la transition verte.

Tableau B.17 - Effets estimés (IV-LSDV)

| Variable | Estimate | Std. Error | t-value | p-value |
|--------------------------------|---|--|--------------------|-----------------------------|
| fit_lag(ΔY_{it}) | 1.0000×10^0 | 1.47×10^{-1} | 6.79×10^1 | $< 2.2 \times 10^{-16} ***$ |
| GreenPotSharew | 4.94×10^{-1} | 2.41×10^{-1} | 2.05 | 0.046* |
| dge | 1.98×10^{-1} | 7.14×10^{-1} | 2.78 | 0.0076** |
| deltalnK | -9.50×10^{-1} | 3.30×10^{-1} | -2.88 | 0.0057** |
| DeltaLnL | 7.45×10^{-1} | 9.47×10^{-1} | 0.79 | 0.43 |
| ays | 7.58×10^{-1} | 6.25×10^{-1} | 1.21 | 0.23 |
| gov_exp_- | 7.05×10^{-1} | 1.91×10^{-1} | 3.69 | 0.00054*** |
| pctgdp | | | | |
| open_pct | -1.18×10^{-1} | 2.14×10^{-1} | -0.55 | 0.58 |
| access_- | 2.10×10^{-1} | 9.61×10^{-1} | 2.18 | 0.034* |
| electricity | | | | |
| GreenPotSharew | -2.996×10^{-1} | 1.482×10^{-1} | -2.02 | 0.048* |
| \times dge | | | | |

Validité de l'instrument et première étape :

- Instrument : lag (ΔY_{it}) est instrumenté par lag (ΔY_{it-1}).
- **Force de l'instrument** : le F-statistique de la première étape est astronomique ($F = 5 \times 10^{32}$, $p < 2 \times 10^{-16}$), ce qui confirme que lag (ΔY_{it}) explique très fortement lag (ΔY_{it-1}).
- **Wu-Hausman** : la statistique de Hausman n'est pas calculée (NA), probablement en raison du très faible bruit dans la première étape ; on suppose malgré tout un biais simultané suffisamment corrigé par l'IV.