

Exercices Module 4

DiD Échelonnée Callaway & Sant'Anna (2021)

Master GPE – UCA FERDI IHEDD

2026-03-24

Objectifs de ce module

À la fin de ces exercices, vous saurez :

1. Utiliser `att_gt()` pour estimer les $ATT(g,t)$
2. Agréger les résultats avec `aggte()`
3. Visualiser avec `ggdid()`
4. Comparer les résultats CS avec le TWFE
5. Interpréter les effets dynamiques et par cohorte

Travailler dans RStudio ?

Téléchargez le script R complet du module avec tous les exercices à compléter :

Ouvrez-le dans RStudio et complétez les `___` au fur et à mesure. Exécutez bloc par bloc avec **Ctrl+Enter**.

Pourquoi ce concept est-il crucial pour un fonctionnaire ?

Dans la pratique des politiques publiques africaines, les programmes sont rarement déployés simultanément partout : une réforme fiscale touche d'abord les grandes villes, puis les villes moyennes, puis les communes rurales. Cette mise en oeuvre *échelonnée* invalide le TWFE classique (module 3) dans certains cas. L'approche Callaway & Sant'Anna répond exactement à ce problème : elle estime des effets séparés pour chaque « cohorte » de bénéficiaires, ce qui vous permet de détecter si l'effet varie selon le moment d'adoption de la réforme – une information précieuse pour les décisions de calendrier de déploiement.

Premiers pas avec `att_gt()`

Exercice 4.1 Estimation de base

Contexte : Imaginez que vous évaluez un programme de modernisation de l'état civil au Bénin. Les communes ont adopté le système biométrique en 3 vagues : 2004 (grandes communes urbaines), 2005 (communes semi-urbaines), 2006 (communes rurales). Le résultat mesuré est le taux d'enregistrement des naissances. Vous utilisez la structure de `mpdta` comme proxy pour cet exemple.

```

{webr-r}
library(did)
library(tidyverse)
data(mpdta)
mpdta$treat_tv <- as.integer(mpdta$first.treat > 0 & mpdta$year >=
mpdta$first.treat)

# TODO : Estimez les ATT(g,t) avec la configuration de base
# Variables :
# - Résultat      : lemp
# - Temps         : year
# - Identifiant   : countyreal
# - Groupe (1er traitement) : first.treat (0 = jamais traité)
# - Groupe de comparaison : "nevertreated"

att_gt_res <- att_gt(
  yname      = "___",      # TODO: variable de résultat
  tname      = "___",      # TODO: variable de temps
  idname     = "___",      # TODO: identifiant des unités
  gname      = "___",      # TODO: année du 1er traitement
  data       = ___,        # TODO: nom du dataset
  control_group = "___",   # TODO: "nevertreated"
  print_details = FALSE
)

summary(att_gt_res)

```

i Indice 1 Si vous êtes bloqué

La fonction `att_gt()` requiert six arguments essentiels : le nom de la variable de résultat (`yname`), le nom de la variable de temps (`tname`), l'identifiant des unités (`idname`), la variable du groupe de traitement (première année de traitement, 0 si jamais traité — `gname`), le dataset (`data`), et le type de groupe de contrôle (`control_group`).

i Indice 2 Indice plus détaillé

```
# Dans mpdta, les variables sont :
# - lemp      → variable de résultat (log emploi)
# - year      → variable temporelle
# - countyreal → identifiant de l'unité (comté)
# - first.treat → année du premier traitement (0 = jamais traité)
# "nevertreated" signifie qu'on compare chaque cohorte aux unités jamais
traitées

att_gt_res <- att_gt(
  yname      = "lemp",
  tname      = "year",
  idname     = "countyreal",
  gname      = "first.treat",
  data       = mpdta,
  control_group = "nevertreated",
  print_details = FALSE
)
```

💡 Voir la solution

```
att_gt_res <- att_gt(
  yname      = "lemp",
  tname      = "year",
  idname     = "countyreal",
  gname      = "first.treat",
  data       = mpdta,
  control_group = "nevertreated",
  print_details = FALSE
)

summary(att_gt_res)
```

Interprétation : `att_gt()` calcule un ATT(g,t) pour chaque combinaison (cohorte g, période t). Pour une cohorte traitée en 2004, vous obtenez des estimations pour chaque année de 2003 à 2007. Les estimations pré-traitement (2003) doivent être proches de zéro. Dans notre contexte béninois fictif, $ATT(g=2004, t=2005)$ nous dirait : « Pour les communes ayant adopté l'état civil biométrique en 2004, quel était l'impact sur le taux d'enregistrement en 2005 ? »

Point clé : Contrairement au TWFE qui produit un seul chiffre, CS produit un tableau d'effets granulaires. C'est plus informatif mais aussi plus complexe à communiquer aux non-économistes. L'agrégation (exercice 4.3) résoudra ce problème.

Exercice 4.2 Visualiser les ATT(g,t)

```
{webr-r}
library(did)
data(mpdta)
```

```

mpdta$treat_tv <- as.integer(mpdta$first.treat > 0 & mpdta$year >=
mpdta$first.treat)

# Recréation de att_gt_res (si nécessaire)
att_gt_res <- att_gt(
  yname = "lemp", tname = "year", idname = "countyreal",
  gname = "first.treat", data = mpdta,
  control_group = "nevertreated", print_details = FALSE
)

# TODO : Visualisez les ATT(g,t) avec ggdid()
# Ajoutez un titre descriptif

ggdid(att_gt_res,
      title = "TODO: Ajoutez un titre") # TODO: complétez le titre

```

💡 Voir la solution et interprétation

```

ggdid(att_gt_res,
      title = "ATT(g,t) – Impact du salaire minimum sur l'emploi des jeunes
(Callaway-Sant'Anna)")

# Interprétation :
# Chaque graphique correspond à un groupe (vague de traitement).
# Les périodes pré-traitement (avant la ligne verticale) doivent avoir des
ATT(g,t)
# proches de zéro.
# Les périodes post-traitement montrent les effets causaux estimés pour chaque
cohorte.

```

Interprétation : Le graphique `ggdid()` produit une grille de panneaux, un par cohorte. Dans chaque panneau, la ligne verticale marque la date de traitement de la cohorte. Les points à gauche de cette ligne sont les pré-tendances ; les points à droite sont les effets causaux. Si les pré-tendances sont bien à zéro dans tous les panneaux, votre identification est solide.

Point clé : L'avantage pédagogique de cette visualisation : elle montre immédiatement si l'effet du programme *varie selon le moment d'adoption*. Dans les politiques publiques, une commune qui adopte une réforme tôt (cohorte 2004) peut avoir un effet différent d'une commune qui l'adopte plus tard (cohorte 2006) — peut-être parce que les adopteurs précoces sont mieux préparés.

Agrégation des résultats

Exercice 4.3 Trois types d'agrégation

```

{webr-r}
library(did)
data(mpdta)
mpdta$treat_tv <- as.integer(mpdta$first.treat > 0 & mpdta$year >=
mpdta$first.treat)

```

```

att_gt_res <- att_gt(
  yname = "lemp", tname = "year", idname = "countyreal",
  gname = "first.treat", data = mpdta,
  control_group = "nevertreated", print_details = FALSE
)

# TODO : Calculez les trois types d'agrégation
# 1. Agrégation simple (une seule valeur : ATT moyen global)
agg_simple <- aggte(att_gt_res, type = "___") # TODO: "simple"


# 2. Agrégation par groupe/cohorte
agg_groupe <- aggte(att_gt_res, type = "___") # TODO: "group"

# 3. Agrégation dynamique (event study)
agg_dynamic <- aggte(att_gt_res, type = "___") # TODO: "dynamic"

# Affichez les résultats
cat("=== ATT moyen global ===\n")
summary(agg_simple)

cat("\n=== ATT par cohorte ===\n")
summary(agg_groupe)

```

 Voir la solution

```

agg_simple <- aggte(att_gt_res, type = "simple")
agg_groupe <- aggte(att_gt_res, type = "group")
agg_dynamic <- aggte(att_gt_res, type = "dynamic")

summary(agg_simple)
summary(agg_groupe)

```

Interprétation : Les trois types d'agrégation répondent à des questions différentes : - "simple" → « Quel est l'effet moyen du programme toutes cohortes confondues ? » (chiffre unique, pour les résumés exécutifs) - "group" → « L'effet diffère-t-il selon la vague d'adoption ? » (pour comprendre si les adopteurs précoces bénéficient plus ou moins) - "dynamic" → « Comment l'effet évolue-t-il dans le temps après l'adoption ? » (pour évaluer si l'effet est immédiat ou s'accumule)

Point clé : Dans un rapport destiné à un ministre, utilisez "simple". Dans une note technique pour des économistes, présentez les trois agrégations. L'agrégation par groupe est particulièrement utile pour comprendre le *sequencing* optimal d'un déploiement progressif.

Exercice 4.4 Event study Callaway-Sant'Anna

```

{webr-r}
library(did)
data(mpdta)
mpdta$treat_tv <- as.integer(mpdta$first.treat > 0 & mpdta$year >=
mpdta$first.treat)

att_gt_res <- att_gt(

```

```

yname = "lemp", tname = "year", idname = "countyreal",
gname = "first.treat", data = mpdta,
control_group = "nevertreated", print_details = FALSE
)

agg_dynamic <- aggte(att_gt_res, type = "dynamic")

# TODO : Visualisez l'event study
# Remplacez les ___ par des arguments appropriés
ggdid(_____, # TODO: objet à visualiser
       title = "Event Study – Callaway & Sant'Anna (2021)",
       ylab = "ATT estimé (log emploi)",
       xlab = "Périodes relatives au traitement")

```

💡 Voir la solution et interprétation

```

ggdid(agg_dynamic,
      title = "Event Study – Callaway & Sant'Anna (2021)",
      ylab = "ATT estimé (log emploi)",
      xlab = "Périodes relatives au traitement")

# Interprétation :
# - k = -1, -2 : pré-tendances (doivent être ≈ 0)
# - k = 0, 1, 2... : effets post-traitement
# L'effet est d'environ -3 à -4% à l'impact, légèrement persistant.
# On conclut que les hausses de salaire minimum ont réduit l'emploi des jeunes
# d'environ 3-4% dans les comtés traités.

```

Interprétation : L'event study agrégé (sur toutes les cohortes) montre la dynamique temporelle de l'effet. Si l'effet augmente en valeur absolue avec le temps ($k=0 < k=1 < k=2$), cela suggère un effet qui s'accumule — typique des programmes d'apprentissage ou de changement de comportement. Si l'effet se stabilise, l'impact est immédiat et permanent.

Point clé : La distinction entre effet *immédiat* et effet *cumulatif* a des implications majeures pour le design des politiques. Un programme dont l'effet s'accumule dans le temps mérite un investissement plus long que prévu initialement.

Référence : Callaway, B. & Sant'Anna, P.H.C. (2021) « Difference-in-Differences with Multiple Time Periods. » *Journal of Econometrics*, 225(2), 200-230.

Comparaison des approches

Exercice 4.5 TWFE vs Callaway-Sant'Anna

```

{webr-r}
library(did)
library(fixest)
library(tidyverse)
data(mpdta)
mpdta$treat_tv <- as.integer(mpdta$first.treat > 0 & mpdta$year >=
mpdta$first.treat)

```

```

# Estimateurs à comparer

# 1. TWFE
twfe_est <- feols(
  lemp ~ treat_tv | countyreal + year,
  data = mpdta,
  cluster = ~countyreal
)

# 2. CS – ATT simple
att_gt_res <- att_gt(
  yname = "lemp", tname = "year", idname = "countyreal",
  gname = "first.treat", data = mpdta,
  control_group = "nevertreated", print_details = FALSE
)
cs_est <- aggte(att_gt_res, type = "simple")

# TODO : Créez un tableau de comparaison
tibble(
  Méthode = c("TWFE (fixest)", "Callaway-Sant'Anna"),
  Estimateur = c(round(coef(___), 4), # TODO: extraire coef TWFE
                 round(cs_est$__ , 4)), # TODO: ATT global CS
  SE = c(round(se(___), 4), # TODO: erreur standard TWFE
         round(cs_est$__ , 4)) # TODO: SE global CS
) |> knitr::kable(caption = "Comparaison TWFE vs Callaway-Sant'Anna")

```

Voir la solution

```

tibble(
  Méthode = c("TWFE (fixest)", "Callaway-Sant'Anna"),
  Estimateur = c(round(coef(twfe_est)["treat_tv"], 4),
                 round(cs_est$overall.att, 4)),
  SE = c(round(se(twfe_est)["treat_tv"], 4),
         round(cs_est$overall.se, 4))
) |> knitr::kable(caption = "Comparaison TWFE vs Callaway-Sant'Anna")

```

```

# Sur ces données, les deux estimateurs sont assez proches.
# Dans des données avec plus d'hétérogénéité, ils peuvent diverger
significativement.

```

Interprétation : Sur mpdta, TWFE et CS donnent des résultats similaires, ce qui indique que l'hétérogénéité des effets entre cohortes est modérée. Dans des situations de déploiement très échelonné avec des effets fortement variables selon la cohorte, les deux estimateurs peuvent diverger substantiellement – et c'est le CS qui est alors plus fiable.

Point clé : Quand TWFE et CS convergent, on a plus confiance dans les résultats. Quand ils divergent, cela est en soi une information importante : l'effet du programme dépend du moment d'adoption, ce qui doit guider la stratégie de déploiement.

Exercice 4.6 Contrôle de l'état de la procédure

```
{webr-r}
library(did)
data(mpdta)
mpdta$treat_tv <- as.integer(mpdta$first.treat > 0 & mpdta$year >=
mpdta$first.treat)

# TODO : Répétez l'estimation en utilisant "notyettreated" au lieu de
"nevertreated"
# et comparez les résultats

att_gt_nyt <- att_gt(
  yname      = "lemp",
  tname      = "year",
  idname     = "countyreal",
  gname      = "first.treat",
  data       = mpdta,
  control_group = "___",          # TODO: "notyettreated"
  print_details = FALSE
)

agg_nyt <- aggte(att_gt_nyt, type = "simple")

# Comparer avec nevertreated
att_gt_nt <- att_gt(
  yname = "lemp", tname = "year", idname = "countyreal",
  gname = "first.treat", data = mpdta,
  control_group = "nevertreated", print_details = FALSE
)

agg_nt <- aggte(att_gt_nt, type = "simple")

cat("=== Comparaison des groupes de contrôle ===\n")
cat("Jamais traités (nevertreated) :", round(agg_nt$overall.att, 4), "\n")
cat("Pas encore traités (notyettreated):", round(agg_nyt$overall.att, 4), "\n")
```

💡 Interprétation

```
# control_group = "notyettreated"

# Interprétation des différences :
# Si les deux estimateurs sont proches, cela renforce la validité des résultats
# (robustesse au choix du groupe de contrôle).
# Si ils divergent, cela peut indiquer que les unités traitées plus tard
diffèrent
# de celles jamais traitées, et que l'hypothèse de tendances parallèles est
plus
# délicate à justifier.
```

Interprétation : Le choix entre "nevertreated" et "notyettreated" est une décision importante. "notyettreated" augmente la taille du groupe de contrôle (on inclut les communes qui seront traitées plus tard comme contrôles temporaires), ce qui donne des estimations plus précises (erreurs standard plus faibles). Mais c'est aussi potentiellement plus sensible si les unités « pas encore traitées » ont des trajectoires différentes des « jamais traitées ».

Point clé : Dans les évaluations de politiques africaines où les groupes de contrôle sont souvent petits (peu de communes jamais traitées), "notyettreated" peut être préférable pour la précision à condition de vérifier la robustesse des résultats entre les deux options.

Discussion Application à votre contexte

Pensez à une politique publique de votre pays ou secteur que vous souhaiteriez évaluer.

1. Quel serait le groupe traité et le groupe de contrôle ?
2. Quelle serait la variable de résultat ?
3. Pourquoi la DiD serait-elle (ou ne serait-elle pas) appropriée ici ?
4. Quelle menace principale à l'hypothèse de tendances parallèles identifieriez-vous ?

Discutez en binôme pendant 5 minutes, puis partagez avec le groupe.

i Fin du Module 4

Compétences validées :

- Estimer les ATT(g,t) avec `att_gt()`
- Agréger avec `aggte()` – simple, groupe, dynamique
- Visualiser avec `ggdid()`
- Comparer TWFE et Callaway-Sant'Anna
- Tester la sensibilité au choix du groupe de contrôle

Prochain exercice : Module 5 – de Chaisemartin & D'Haultfoeuille avec `DIDmultiplegtDYN`