

Économétrie causale appliquée

Informations generales

Element	Detail
Intitulé	Économétrie causale appliquée
Niveau	Master 1 / Master 2 — TD + Projet R
Volume horaire	24h TD + Projet (12 séances de 2h)
Crédits	4 ECTS
Semestre	S1 ou S2
Enseignant	Eric Gabin Kilama
Affiliation	FERDI · CERDI
Contact	kilamaericgabin@yahoo.fr
Logiciels	R (fixest, rdrobust, sandwich), Stata, Python (pandas, statsmodels)
Langue	Français (supports et lectures en anglais)

Prerequis

- **Statistique L3** : distributions, tests d'hypotheses, intervalles de confiance, vraisemblance.
 - **Econometrie L3** : modele lineaire, MCO, proprietes des estimateurs, heteroscedasticite, tests de specification.
 - **Logiciel** : familiarite avec au moins un logiciel statistique (Stata, R ou Python). Les TP proposent des scripts dans les trois environnements.
-

Description du cours

Ce cours forme les etudiants de Master 1 aux methodes econometriques modernes d'identification causale, en insistant sur la rigueur de la mise en oeuvre empirique et la reproductibilite des resultats. A partir de donnees reelles issues de questions de politique publique (commerce international, aide au developpement, sanctions economiques, gouvernance), les etudiants apprennent a diagnostiquer les biais d'un modele lineaire, a choisir et implementer la strategie d'identification appropriee (variables instrumentales, panels, regression sur discontinuite, differences de differences), et a communiquer leurs resultats de maniere transparente. Le cours integre les avancees recentes de la litterature methodologique, notamment les corrections pour le DiD echelonne (Sun et Abraham, 2021 ; Callaway et Sant'Anna, 2021), les methodes de selection de variables en haute dimension (lasso, post-lasso), et les standards de reproductibilite (gestion de code, documentation, replication packages). Chaque seance articule theorie, implementation informatique et discussion critique d'articles empiriques publies dans des revues de premier plan.

Objectifs pedagogiques

A l'issue du cours, les etudiants seront capables de :

1. **Diagnostiquer** les problemes d'un modele de regression lineaire (endogeneite, heteroscedasticite, autocorrelation, multicolinearite) et choisir les corrections appropriees.
 2. **Implementer** les principales strategies d'identification causale (IV/2SLS, panel FE/RE, RDD, DiD classique et echelonne) dans Stata, R ou Python.
 3. **Evaluer** la validite des hypotheses d'identification dans un article empirique publie et formuler des critiques constructives.
 4. **Maitriser** les outils de reproductibilite (structure de projet, controle de version, documentation du code, set.seed) et produire un replication package complet.
 5. **Interpreter** les resultats econometriques dans leur contexte de politique publique, en distinguant significativite statistique, significativite economique et causalite.
 6. **Presenter** une analyse empirique de maniere professionnelle, a l'oral comme a l'ecrit, en respectant les standards des revues academiques.
-

Manuels et ressources de reference

Manuels principaux

- **Angrist, J.D. et J.-S. Pischke** (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. — Reference centrale du cours pour l'identification causale.
- **Wooldridge, J.M.** (2010). *Econometric Analysis of Cross Section and Panel Data*. 2e edition, MIT Press. — Reference pour les methodes de panel et les modeles non lineaires.

Manuels complementaires

- **Cattaneo, M.D., N.S. Idrobo et R. Titiunik** (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge University Press. — Reference pour les seances RDD.
- **Cunningham, S.** (2021). *Causal Inference: The Mixtape*. Yale University Press. — Complement accessible avec exemples codes.
- **Huntington-Klein, N.** (2022). *The Effect: An Introduction to Research Design and Causality*. Chapman & Hall. — Excellent pour l'intuition graphique.

Ressources logicielles

- **Stata** : documentation officielle + help integre. Package reghdfe (Correia, 2016), ivreghdfe, rdrobust.
 - **R** : package fixest (Berge, 2018) pour les estimations a effets fixes haute dimension, rdrobust (Cattaneo et al.), sandwich (Zeileis et al.), did (Callaway et Sant'Anna).
 - **Python** : linearmodels (panel), statsmodels (OLS, IV), rdrobust (port Python).
-

Programme des seances

Seance 1 – MCO : Diagnostics et limites

Objectifs : Revoir les proprietes des MCO, identifier les violations des hypotheses classiques, maitriser les tests de specification.

CM (1h30)

1. Rappel des hypotheses du modele lineaire classique (Gauss-Markov)
2. Diagnostic de l'heteroscedasticite : tests de Breusch-Pagan, White
3. Diagnostic de la multicollinearite : VIF, condition number
4. Diagnostic de la forme fonctionnelle : test RESET de Ramsey
5. Consequences des violations : biais, inefficacite, inference invalide
6. Introduction au biais de variable omise : formalisation et exemples

Readings obligatoires

- Angrist, J.D. et J.-S. Pischke (2009). *Mostly Harmless Econometrics*, chapitre 3 (“Making regression make sense”).
- Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data*, chapitre 4 (“The single-equation linear model and OLS estimation”).

Readings complementaires

- Goldberger, A.S. (1991). *A Course in Econometrics*. Harvard University Press, chapitres 14-16.

TP (1h15)

Exercice : Rendements de l'education et biais de variable omise. A partir des donnees CPS (Current Population Survey), estimer une equation de Mincer par MCO. Diagnostiquer l'heteroscedasticite (test de White), la forme fonctionnelle (RESET), calculer les VIF. Comparer les resultats avec et sans controles pour l'aptitude (score AFQT). Interpreter le changement de coefficient comme evidence du biais de variable omise. Logiciels : Stata (reg, estat hetttest, estat vif, ovtest) et R (lm, bptest, vif).

Seance 2 – Erreurs standards robustes et clustering

Objectifs : Comprendre pourquoi les erreurs standards classiques sont souvent invalides et maitriser les corrections modernes.

CM (1h30)

1. Erreurs standards robustes a l'heteroscedasticite (White/Eicker-Huber-White)
2. Erreurs standards clusterisees : intuition, formule, choix du niveau de clustering
3. Le probleme des “few clusters” : corrections (wild cluster bootstrap, CR2/CR3)
4. Erreurs standards HAC (Newey-West) pour les series temporelles
5. Regles pratiques : quand clusterer, a quel niveau, combien de clusters minimum

6. Exemples : clustering par pays dans les regressions de gravite, clustering par etat dans les DiD

Readings obligatoires

- Cameron, A.C. et D.L. Miller (2015). “A Practitioner’s Guide to Cluster-Robust Inference”. *Journal of Human Resources*, 50(2), 317-372.
- Angrist, J.D. et J.-S. Pischke (2009). *Mostly Harmless Econometrics*, chapitre 8 (“Nonstandard standard error issues”).

Readings complementaires

- MacKinnon, J.G., M.O. Nielsen et M.D. Webb (2023). “Cluster-Robust Inference: A Guide to Empirical Practice”. *Journal of Econometrics*, 232(2), 272-299.
- Abadie, A., S. Athey, G.W. Imbens et J. Wooldridge (2023). “When Should You Adjust Standard Errors for Clustering?” *Quarterly Journal of Economics*, 138(1), 1-35.

TP (1h15)

Exercice : Commerce bilateral et choix de clustering. A partir d’un extrait de donnees de gravite (CEPII), estimer une equation de gravite log-lineaire. Comparer les erreurs standards : homoskedastiques, robustes HC1, clusterisees par paire exportateur-importateur, clusterisees par exportateur, double-clusterisees (exportateur + importateur). Observer comment la significativite change selon le choix. Implementer le wild cluster bootstrap avec le package `boottest` (Stata) ou `fwildclusterboot` (R). Logiciels : Stata (`reghdfe`, `boottest`) et R (`fixest::feols` avec `vcov`).

Seance 3 – Variables instrumentales

Objectifs : Maitriser l’identification par variables instrumentales, de la theorie a l’implementation, en evaluant la validite des instruments.

CM (1h30)

1. Le probleme d’endogeneite : simultaneite, erreur de mesure, variable omise
2. Conditions de validite d’un instrument : pertinence (first stage) et exclusion
3. Estimation 2SLS : derivation, proprietes, interpretation
4. Diagnostics : test de sous-identification (Kleibergen-Paap), instruments faibles (F de Cragg-Donald, F effectif de Olea-Pflueger), test de suridentification (Hansen J)
5. Interpretation LATE vs ATE : le theoreme de Wald et les compliers
6. Exemples classiques : precipitations et conflits (Miguel et al., 2004), institutions coloniales et developpement (Acemoglu et al., 2001)

Readings obligatoires

- Angrist, J.D. et J.-S. Pischke (2009). *Mostly Harmless Econometrics*, chapitre 4 (“Instrumental variables in action”).
- Stock, J.H. et M. Yogo (2005). “Testing for Weak Instruments in Linear IV Regression”. In *Identification and Inference for Econometric Models*, Cambridge University Press, 80-108.

Readings complementaires

- Andrews, I., J.H. Stock et L. Sun (2019). “Weak Instruments in Instrumental Variables Regression: Theory and Practice”. *Annual Review of Economics*, 11, 727-753.
- Lee, D.S., J. McCrary, M.J. Moreira et J. Porter (2022). “Valid t-ratio Inference for IV”. *American Economic Review*, 112(10), 3260-3290.

TP (1h15)

Exercice : Institutions et développement. Replier partiellement Acemoglu, Johnson et Robinson (2001) : estimer l’effet des institutions (indice de protection contre l’expropriation) sur le PIB par tete, en instrumentant par la mortalite des colons. Estimer le first stage, calculer le F de Cragg-Donald, tester la suridentification avec un second instrument (latitude). Comparer MCO et 2SLS. Discuter la restriction d’exclusion. Logiciels : Stata (ivregress 2sls, estat firststage, estat overid) et R (fixest::feols avec syntaxe IV, summary diagnostics).

Seance 4 – Donnees de panel : effets fixes et effets aleatoires

Objectifs : Exploiter la dimension longitudinale des donnees pour controler l’heterogeneite inobservee et choisir entre FE et RE.

CM (1h30)

1. Structure des donnees de panel : notation, balanced vs unbalanced
2. Modele a effets fixes (within) : elimination de l’heterogeneite individuelle, estimation, interpretation
3. Modele a effets aleatoires (GLS) : hypotheses, estimation, gains d’efficience
4. Test de Hausman : FE vs RE, interpretation, limites
5. Effets fixes haute dimension : multiples niveaux de FE, estimation par absorption (algorithme d’Inoue-Solon, methode de Correia)
6. Panel dynamique : biais de Nickell, GMM d’Arellano-Bond (introduction)
7. Exemples : determinants de la croissance, effet des accords commerciaux, impact de l’aide

Readings obligatoires

- Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data*, chapitres 10-11 (“Basic linear unobserved effects panel data models” et “More topics in linear unobserved effects models”).
- Correia, S. (2016). “Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator”. Working Paper. — Documentation technique de reghdfe.

Readings complementaires

- Mundlak, Y. (1978). “On the Pooling of Time Series and Cross Section Data”. *Econometrica*, 46(1), 69-85.
- Berge, L. (2018). “Efficient Estimation of Maximum Likelihood Models with Multiple Fixed-Effects: the R Package fixest”. CREA Discussion Paper. — Documentation de fixest.

TP (1h15)

Exercice : Aide internationale et croissance. A partir d’un panel de pays en developpement (Penn World Table + donnees d’aide OCDE-CAD), estimer l’effet de l’aide sur la croissance du PIB par

tete. Comparer pooled OLS, RE et FE. Realiser le test de Hausman. Ajouter des effets fixes temporels. Estimer avec des effets fixes haute dimension (pays + annee + region x annee). Observer la stabilite des coefficients. Logiciels : Stata (xtreg, reghdfe) et R (fixest::feols avec notation | country + year + region^year).

Seance 5 – Regression sur discontinuite (RDD)

Objectifs : Maitriser le design RDD (sharp et fuzzy), les choix de bande passante et les tests de validite.

CM (1h30)

1. Intuition de la RDD : exploitation d'un seuil exogene
2. RDD sharp : estimation non parametrique (polynomes locaux), choix de bande passante (MSE-optimal de Calonico, Cattaneo et Titiunik, 2014)
3. RDD fuzzy : analogie avec l'IV, estimation en deux etapes
4. Tests de validite : test de manipulation de la running variable (McCrary, 2008 ; Cattaneo, Jansson et Ma, 2020), tests de continuite des covariables
5. Extensions : RDD géographique, RDD multi-seuils, RDD avec covariables
6. Exemples : effets electoraux (Lee, 2008), aide internationale et seuils de revenu (Galiani et al., 2017)

Readings obligatoires

- Cattaneo, M.D., N.S. Idrobo et R. Titiunik (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*, chapitres 1-4.
- Lee, D.S. et T. Lemieux (2010). "Regression Discontinuity Designs in Economics". *Journal of Economic Literature*, 48(2), 281-355.

Readings complementaires

- Calonico, S., M.D. Cattaneo et R. Titiunik (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". *Econometrica*, 82(6), 2295-2326.
- Cattaneo, M.D., R. Titiunik et G. Vazquez-Bare (2020). "The Regression Discontinuity Design". In *Handbook of Research Methods in Political Science and International Relations*, Sage, 835-857.

TP (1h15)

Exercice : *Seuils d'eligibilite et aide internationale*. A partir de donnees simulees reproduisant la logique de Galiani et al. (2017) — seuil de revenu par tete pour l'eligibilite IDA de la Banque mondiale —, estimer l'effet de l'aide sur un indicateur de developpement. Implementer la RDD sharp avec polynomes locaux et bande passante MSE-optimale. Realiser le test de McCrary sur la running variable. Tester la continuite des covariables au seuil. Comparer estimations avec differentes bandes passantes et ordres de polynome. Logiciels : Stata (rdrobust, rddensity) et R (rdrobust::rdrobust, rddensity::rddensity).

Seance 6 – Differences de differences classique et echelonne

Objectifs : Maitriser le DiD a deux periodes et comprendre les biais du TWFE avec traitement echelonne, puis implementer les correctifs modernes.

CM (1h30)

1. DiD classique a deux periodes et deux groupes : hypothese de tendances paralleles, estimation, tests de pre-tendance
2. TWFE generalise : notation, hypotheses, interpretation des coefficients
3. Le probleme du traitement echelonne : biais de TWFE avec effets heterogenes (Goodman-Bacon, 2021 ; de Chaisemartin et D’Haultfoeuille, 2020)
4. Decomposition de Goodman-Bacon : quels groupes pesent dans l’estimateur TWFE
5. Estimateurs robustes : Callaway et Sant’Anna (2021), Sun et Abraham (2021), Borusyak, Jaravel et Spiess (2024), de Chaisemartin et D’Haultfoeuille (2020)
6. Choix pratique entre estimateurs : arbre de decision
7. Exemples : effets des sanctions economiques, reformes de politique commerciale

Readings obligatoires

- Goodman-Bacon, A. (2021). “Difference-in-Differences with Variation in Treatment Timing”. *Journal of Econometrics*, 225(2), 254-277.
- Callaway, B. et P.H.C. Sant’Anna (2021). “Difference-in-Differences with Multiple Time Periods”. *Journal of Econometrics*, 225(2), 200-230.

Readings complementaires

- Sun, L. et S. Abraham (2021). “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects”. *Journal of Econometrics*, 225(2), 175-199.
- Roth, J., P.H.C. Sant’Anna, A. Bilinski et J. Poe (2023). “What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature”. *Journal of Econometrics*, 235(2), 2218-2244.

TP (1h15)

Exercice : Reformes commerciales et performance a l’exportation. A partir d’un panel simule de pays avec adoption echelonnee d’accords de libre-echange, estimer l’effet sur les exportations. D’abord en TWFE classique (reghdfe/fixest). Puis realiser la decomposition de Goodman-Bacon pour identifier les comparaisons problematiques. Implementer l’estimateur de Callaway-Sant’Anna et l’estimateur de Sun-Abraham. Comparer les event-study plots. Observer les divergences et interpreter. Logiciels : Stata (did_multiplegt, csdid, eventstudyinteract) et R (did::att_gt, fixest::sunab).

Seance 7 – Haute dimension et selection de variables

Objectifs : Introduire les methodes de regularisation (lasso, post-lasso, double selection) pour les contextes ou le nombre de controles potentiels est eleve.

CM (1h30)

1. Le probleme de la haute dimension : beaucoup de controles potentiels, risque de surapprentissage

2. Regularisation L1 (lasso) : intuition geometrique, choix du parametre de penalisation (validation croisee, BIC)
3. Post-lasso : estimation OLS sur les variables selectionnees
4. Double selection de Belloni, Chernozhukov et Hansen (2014) : pourquoi le lasso seul ne suffit pas pour l'inference causale
5. Double/debiased machine learning (Chernozhukov et al., 2018) : cadre general, cross-fitting
6. Applications en economie : selection de controles dans les equations de salaire, prediction de la croissance
7. Limites : lasso n'identifie pas les causes, selection vs prediction

Readings obligatoires

- Belloni, A., V. Chernozhukov et C. Hansen (2014). "Inference on Treatment Effects after Selection among High-Dimensional Controls". *Review of Economic Studies*, 81(2), 608-650.
- Mullainathan, S. et J. Sperber (2017). "Machine Learning: An Applied Econometric Approach". *Journal of Economic Perspectives*, 31(2), 87-106.

Readings complementaires

- Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey et J. Robins (2018). "Double/Debiased Machine Learning for Treatment and Structural Parameters". *Econometrics Journal*, 21(1), C1-C68.
- Athey, S. et G.W. Imbens (2019). "Machine Learning Methods That Economists Should Know About". *Annual Review of Economics*, 11, 685-725.

TP (1h15)

Exercice : Selection de determinants de la croissance. A partir d'un dataset de croissance a la Barro (panel de pays, ~60 variables explicatives potentielles), implementer le lasso, le post-lasso et la double selection. Comparer les variables retenues par chaque methode. Estimer l'effet d'une variable d'interet (ouverture commerciale) en controlant par les variables selectionnees. Comparer avec un MCO "kitchen sink". Discuter les implications pour l'inference. Logiciels : Stata (lasso2, pds lasso) et R (glmnet, hdm, DoubleML).

Seance 8 – Reproductibilite et workflow empirique

Objectifs : Maitriser les standards de reproductibilite en economie empirique et structurer un projet de recherche complet.

CM (1h30)

1. La crise de reproductibilite en sciences sociales : resultats non replicables, p-hacking, garden of forking paths
2. Standards actuels des revues : data availability policies (AER, QJE, REStud), replication packages
3. Structure d'un projet reproductible : arborescence (data/raw, data/clean, code, output), README, licence
4. Controle de version avec git : commits, branches, collaboration
5. Automatisation : Makefiles, scripts master, documentation du pipeline donnees → resultats
6. Bonnes pratiques de codage : set.seed(), chemins relatifs, commentaires, logs

7. Ethique des données : anonymisation, consentement, accès restreint

Readings obligatoires

- Christensen, G. et E. Miguel (2018). “Transparency, Reproducibility, and the Credibility of Economics Research”. *Journal of Economic Literature*, 56(3), 920-980.
- Gentzkow, M. et J.M. Shapiro (2014). “Code and Data for the Social Sciences: A Practitioner’s Guide”. Mimeo, University of Chicago.

Readings complémentaires

- Vilhuber, L. (2020). “Reproducibility and Replicability in Economics”. *Harvard Data Science Review*, 2(4).
- Orozco, V., C. Bontemps, E. Maigne, V. Orozco-Olvera, A. Rossi et C. Salanié (2020). “How to Make a Pie: Reproducible Research for Empirical Economics and Econometrics”. *Journal of Economic Surveys*, 34(5), 1134-1169.

TP (1h15)

Exercice : Construire un projet reproductible de A à Z. À partir d’un dataset fourni et d’un script d’estimation non structure, les étudiants doivent : (1) créer l’arborescence de projet, (2) séparer nettoyage et estimation en scripts distincts, (3) écrire un script master qui exécute le pipeline complet, (4) initialiser un dépôt git et faire 3 commits logiques, (5) documenter dans un README. Évaluation sur la capacité d’un pair à reproduire les résultats à partir du dépôt seul. Logiciels : terminal (git, make), éditeur de code.

Seance 9 – Projet de replication : atelier

Objectifs : Appliquer l’ensemble des méthodes du cours à la replication d’un article empirique publiée, en identifiant les forces et faiblesses de l’identification.

CM (1h30)

1. Qu’est-ce qu’une replication ? Replication exacte, replication avec extension, analyse de robustesse
2. Comment lire un article empirique : identification, données, résultats, limites
3. Grille d’évaluation d’une stratégie d’identification : pertinence, exclusion, tendances parallèles, manipulation
4. Comment rédiger un rapport de replication : structure, ton, contribution
5. Exemples de replications publiées : *Journal of Applied Econometrics Data Archive*, I4R (Institute for Replication)
6. Présentation des articles proposés pour le projet final

Readings obligatoires

- Hamermesh, D.S. (2007). “Replication in Economics”. *Canadian Journal of Economics*, 40(3), 715-733.
- Clemens, M.A. (2017). “The Meaning of Failed Replications: A Review and Proposal”. *Journal of Economic Surveys*, 31(1), 326-342.

Readings complémentaires

- Brodeur, A., N. Cook et A. Heyes (2020). “Methods Matter: p-Hacking and Publication Bias in Causal Analysis in Economics”. *American Economic Review*, 110(11), 3634-3660.

TP (1h15)

Atelier guide : démarrage du projet de replication. En groupes de 2-3, les étudiants choisissent un article parmi la liste proposée (articles avec données et code publiquement disponibles). Pendant la séance : (1) télécharger les données et le code, (2) exécuter le code original et vérifier la reproductibilité exacte, (3) identifier une table ou figure clé à répliquer en détail, (4) lister les extensions possibles (échantillon différent, méthode alternative, test de robustesse additionnel). Chaque groupe présente son plan en 3 minutes.

Seance 10 – Presentations étudiants et synthèse

Objectifs : Présenter les projets de replication, recevoir des retours critiques des pairs et de l’enseignant, synthétiser les méthodes du cours.

CM (0h45)

1. Synthèse du cours : l’arbre de décision méthodologique
 - Biais de variable omise □ FE ou IV
 - Seuil exogène □ RDD
 - Choc temporel □ DiD (staggered si adoption progressive)
 - Beaucoup de contrôles □ lasso/double selection
2. Frontières de la recherche : synthetic control, shift-share (Bartik), bunching, méthodes ML causales
3. Conseils pour le mémoire de M2 et la thèse

Presentations étudiants (2h)

Chaque groupe (2-3 étudiants) présente son projet de replication en 12 minutes + 5 minutes de questions. Structure attendue : (1) article original et question de recherche, (2) stratégie d’identification et hypothèses clés, (3) résultats de la replication exacte, (4) extension ou robustesse additionnelle, (5) évaluation critique de l’identification. Évaluation par l’enseignant et les pairs (grille fournie).

Readings obligatoires

- Aucune lecture supplémentaire. Relire les notes de cours et préparer la présentation.

Readings complémentaires

- Angrist, J.D. et J.-S. Pischke (2010). “The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics”. *Journal of Economic Perspectives*, 24(2), 3-30.

Evaluation

Element	Poids	Description
Participation TP	20%	Presence, engagement, qualite du code soumis a chaque seance
Mini-analyses	30%	3 devoirs individuels (seances 3, 5, 7) : application d'une methode a un dataset fourni, rapport de 3-4 pages
Projet de replication	50%	En groupe (2-3). Replication + extension d'un article publie. Livrable : replication package complet (code, donnees, README) + rapport de 10-15 pages + presentation orale (seance 10)

Criteres du projet de replication

- Replication exacte des resultats principaux (15/50)
- Extension ou robustesse additionnelle (15/50)
- Qualite du code et de la documentation (10/50)
- Qualite de la critique methodologique (10/50)

Lien recherche-enseignement

Ce cours s'appuie directement sur l'experience de recherche de l'enseignant en economie appliquee. Les methodes enseignees sont celles utilisees dans des travaux publies ou en cours :

Methode du cours	Lien avec la recherche	Application
Meta-regression (MCO, WLS)	Article publie dans l' <i>European Economic Review</i>	Evaluation de l'effet du ciblage d'inflation dans les pays en developpement par meta-analyse
Panel IV, effets fixes	Article publie dans le <i>Journal of Comparative Economics</i>	Analyse de l'impact de la corruption sur l'investissement, identification par panel et instruments
Panel IV haute dimension	Article en revision a l' <i>AEJ: Economic Policy</i>	Effet des coups d'Etat sur la corruption, instrumentation par contagion regionale
DiD echelonne	Travaux en cours sur les PME et risques geopolitiques	Application des estimateurs de Callaway-Sant'Anna aux chocs geopolitiques

Methode du cours	Lien avec la recherche	Application
PPML, gravite	Travaux en cours sur le commerce et les conflits	Estimation structurelle des effets des sanctions sur les flux commerciaux
Reproductibilite	Experience au CAPS/MEAE	Pratique de la production d'analyses reproductibles pour l'aide a la decision publique

Positionnement pedagogique

Ce cours se distingue des enseignements standards d'econometrie de M1 sur trois dimensions :

1. **Methodes modernes** : au-dela du TWFE classique, le cours couvre les avancees recentes sur le DiD echelonne (Sun-Abraham, Callaway-Sant'Anna), la haute dimension (double selection, lasso), et les corrections d'inference (wild cluster bootstrap). Ces methodes sont desormais standard dans les meilleures revues mais rarement enseignees en M1.
2. **Reproductibilite integree** : la seance 8 et le projet final imposent les standards de reproductibilite (git, arborescence, documentation) comme competence a part entiere, pas comme option. Les etudiants sortent du cours avec l'habitude de structurer leurs projets empiriques de maniere transparente et replicable.
3. **Ancrage policy** : les exemples et exercices mobilisent des questions de politique publique reelles — sanctions economiques, aide au developpement, accords commerciaux, gouvernance — plutot que des exemples generiques. Cette orientation reflete l'experience de l'enseignant, qui a pratique l'analyse econometrique aussi bien en contexte academique (publications en revues) qu'en contexte institutionnel (experience au CAPS/MEAE). Les etudiants apprennent a lire les coefficients non seulement comme des parametres statistiques mais comme des elements de diagnostic pour la decision publique.