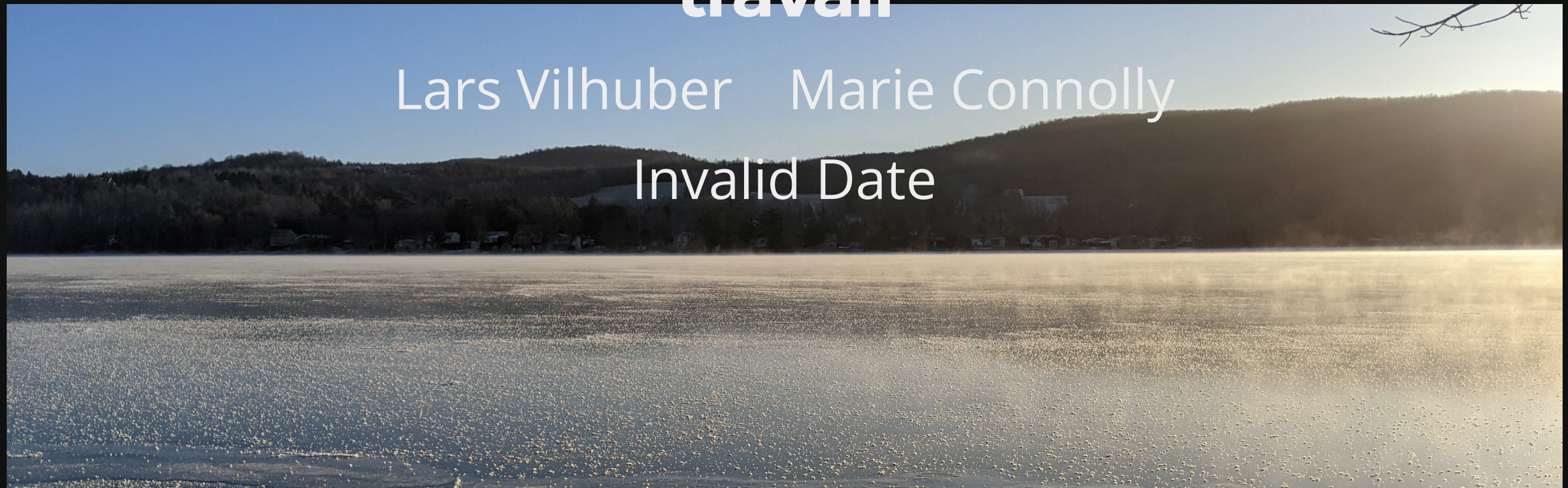
Reproductibilité en recherche – Assurer la transparence et la crédibilité de votre travail

Lars Vilhuber Marie Connolly

Invalid Date



Suivez en direct

larsvilhuber.github.io/ecole-d-ete-ciqss-2026/presentation/presentation.html (PDF)

Qui sommes-nous ?

Vos instructeurs

Marie Connolly, UQAM

Lars Vilhuber, Cornell University

Marie Connolly

Professeure titulaire, Sciences économiques, École des sciences de la gestion (ESG), UQAM. Sa recherche porte sur divers sujets en économie du travail, tels que la mobilité intergénérationnelle des revenus, la formation du capital humain, l'écart entre les sexes et la famille, la participation des femmes au marché du travail et l'évaluation des politiques publiques. Elle est la "Data Editor" de la Revue canadienne d'économique.

Lars Vilhuber

American Economic Review

The *American Economic Review* is a general-interest economics journal. Established in 1911, the *AER* is among the nation's oldest and most respected scholarly journals in economics.

American Economic Review: Insights

AER: Insights is designed to be a top-tier, general-interest economics journal publishing papers of the same quality and importance as those in the *AER*, but devoted to publishing papers with important insights that can be conveyed succinctly.

Journal of Economic Literature

The *Journal of Economic Literature (JEL)*, first published in 1969, is designed to help economists keep abreast of and synthesize the vast flow of literature.

Journal of Economic Perspectives

The *Journal of Economic Perspectives (JEP)* fills the gap between the general interest press and academic economics journals.

American Economic Journal: Applied Economics

American Economic Journal: Applied Economics publishes papers covering a range of topics in applied economics, with a focus on empirical microeconomic issues.

American Economic Journal: Economic Policy

American Economic Journal: Economic Policy publishes papers covering a range of topics, the common theme being the role of economic policy in economic outcomes.

American Economic Journal: Macroeconomics

American Economic Journal: Macroeconomics focuses on studies of aggregate fluctuations and growth, and the role of policy in that context.

American Economic Journal: Microeconomics

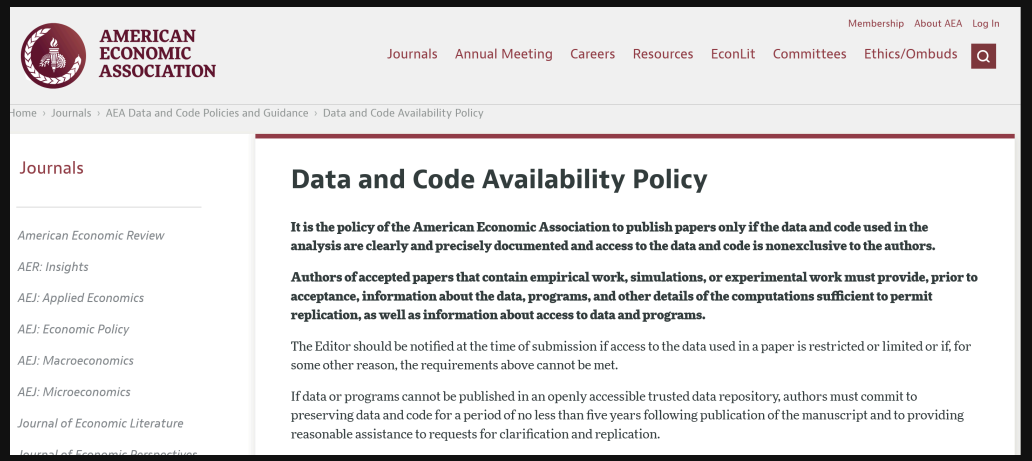
American Economic Journal: Microeconomics publishes papers focusing on microeconomic theory; industrial organization; and the microeconomic aspects of international trade, political economy, and finance.

Directeur exécutif du Labor Dynamics Institute et chercheur associé principal au Département d'économie de Cornell University, et Data Editor de l'American Economic Association.

Revues

Éditeur de données de l'AEA

2791 manuscrits et 4470 rapports, environ 5000 auteurs contactés.



DCAP

Quelques points pratiques

Langues

- Toutes les diapositives sont en **français**
- La langue principale est le **français**
- All questions can be in ***English***
- All answers will be in the ***language asked***, or in French

Code de conduite

- voir Code de conduite complet du LDI Replication Lab, Association canadienne d'économique (en français)

> Nous nous engageons à offrir un environnement accueillant et de soutien pour toutes les personnes, indépendamment de leur origine ou de leur identité. En participant à cette équipe, les participants acceptent de respecter le Code de conduite du LDI ReplicationLab et d'accepter les procédures de résolution de tout incident lié au Code de conduite. Toute forme de comportement visant à exclure, intimider ou causer de l'inconfort constitue une violation du Code de conduite. Afin de favoriser un environnement d'apprentissage positif et professionnel, nous encourageons les types de comportements suivants sur toutes les plateformes et lors de tous les événements :

Code de conduite

- Utiliser un langage **accueillant** et inclusif
- Être ***respectueux*** des différents points de vue et expériences
- Accepter gracieusement les critiques **constructives**
- Faire preuve de courtoisie et de respect envers les autres membres de la communauté

Si vous pensez que quelqu'un viole le Code de conduite, nous vous demandons de le signaler à nous, au CIQSS (commanditaire) ou à l'ACÉ (hôte).

Walkthrough of the agenda

Today

- 9h00 Bienvenue
- 9h05 Tour d’horizon de l’ordre du jour
- 9h15 Objectifs
- 9h30 Configuration technique, formation possible d’équipes
- 9h45 🔒 *Exercice pratique : Un exemple très imparfait*
- 10h00 **Day 1** : Se préparer à la reproductibilité
- 11h30 Tout documenter : Comment documenter correctement un paquet de réplication (et pourquoi !)

Tomorrow

- 13h00 🔒Quand le nettoyage de données est 🔒 critique
- 13h30 **Sujet B** (voir le sondage)
- 14h15 Pause
- 14h30 ***Pratique :*** Améliorer le paquet de réplication (très imparfait -> beaucoup mieux)
- 15h00 ***Pratique :*** Tout tester
- 15h15 Conclusion
- 16h00 Fin.

Meilleures pratiques

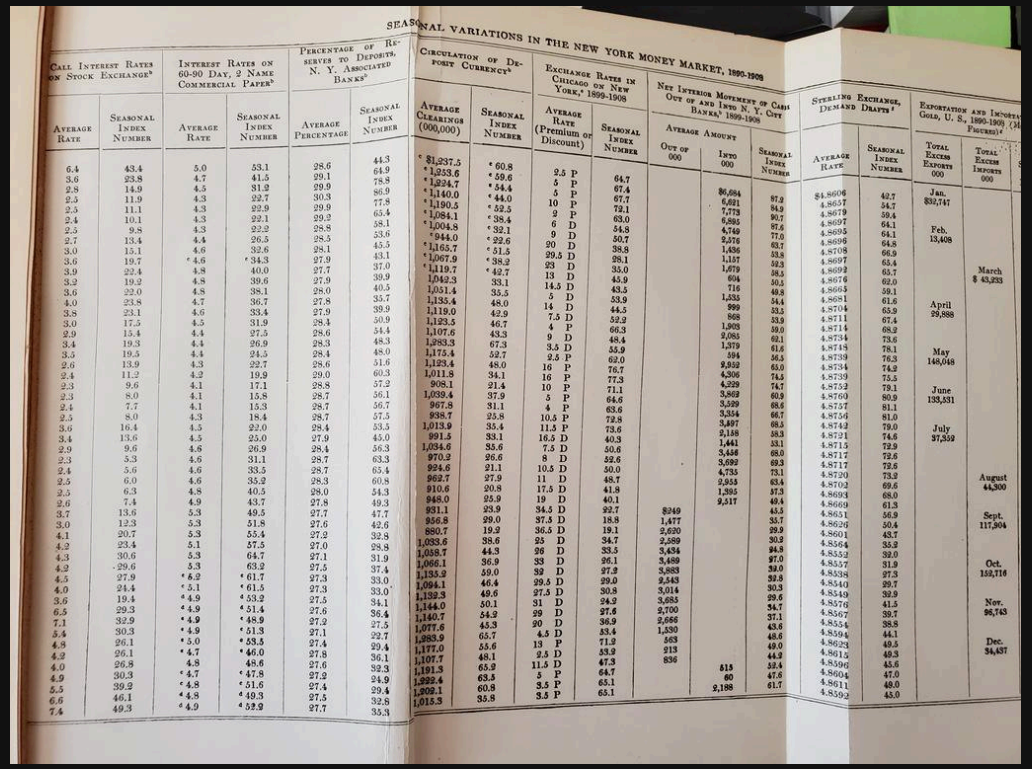

D’abord : pourquoi ?

Pourquoi la reproductibilité ?

- Crédibilité
- Transparence (ouverture)
- Efficacité du discours scientifique ?

Pourquoi la reproductibilité ?

- Les premières publications (20e siècle) ***contenaient des tableaux de données***, et les mathématiques étaient simples (peut-être)
- Les données sont devenues électroniques, **n'étaient plus incluses** ou citées
- Les mathématiques ont été transcrites en ***code***, et n'étaient ***plus incluses***



SEASONAL VARIATIONS IN THE NEW YORK MONEY MARKET, 1890-1908

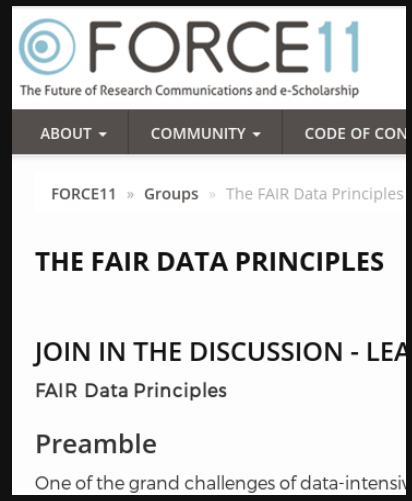
Call Interest Rates on Stock Exchange[b]		Interest Rates on 60-90 Day, 2 Name Commercial Paper[b]		Percentage of Reserves to Deposits N. Y. Associated Banks[b]		Circulation of Deposit Currency[b]		Exchange Rates in Chicago on New York,[c] 1899-1908		Net Interior Movement of Cash Out of and Into N. Y. City Banks,[b] 1899-1908			Sterling Exchange, Demand Drafts[c]		Exportation and Importation Gold, U. S., 1890-1909 (34 Figures)[d]	
Average Rate	Seasonal Index Number	Average Rate	Seasonal Index Number	Average Percentage	Seasonal Index Number	Average Clearings (000,000)	Seasonal Index Number	Average Rate (Premium or Discount)	Seasonal Index Number	Average Amount Out of 000	Average Amount Into 000	Seasonal Index Number	Average Rate	Seasonal Index Number	Total Excess Exports 000	Total Excess Imports 000
6.4	43.4	5.0	53.1	28.6	44.3	[c]$1,237.5	[c]60.8									
3.6	25.8	4.7	41.5	29.1	64.9	[c]1,253.6	[c]59.6	2.5 P	64.7							
2.8	14.9	4.5	31.2	29.9	78.8	[c]1,234.7	[c]54.4	5 P	67.4							
2.5	11.9	4.3	22.7	30.3	86.9	[c]1,140.0	[c]44.0	5 P	67.7							
2.5	11.1	4.3	22.9	29.9	77.8	[c]1,190.5	[c]52.5	10 P	72.1							
2.4	10.1	4.3	22.1	29.2	65.4	[c]1,084.1	[c]38.4	2 P	63.0							
2.5	9.8	4.3	22.2	28.8	58.1	[c]1004.8	[c]32.1	6 D	54.8							
2.7	13.4	4.4	26.5	28.5	53.6	[c]944.0	[c]22.6	9 D	50.7							
3.0	15.1	4.6	32.6	28.1	45.5	[c]1165.7	[c]51.5	20 D	38.8							
3.6	19.7	[c]4.6	[c]34.3	27.9	43.1	[c]1067.9	[c]38.2	23 D	28.1							
3.9	22.4	4.8	40.0	27.7	37.0	[c]1119.7	[c]42.7	13 D	33.0							
3.2	19.2	4.8	39.6	27.9	39.9	1,042.3	33.1	14.5 D	45.9							
3.6	22.0	4.8	38.1	28.0	40.5	1,051.4	35.5	5 D	45.5							
4.0	23.8	4.7	36.7	27.8	35.7	1,135.4	48.0	14 D	53.9							
3.8	23.1	4.6	33.4	27.9	39.9	1,119.0	42.9	7.5 D	44.5							
3.0	17.5	4.5	31.9	28.4	50.9	1,123.5	46.7	4 P	52.2							
2.9	15.4	4.4	27.5	28.6	54.4	1,107.6	43.3	9 D	66.3							
3.4	19.3	4.4	26.9	28.3	48.3	1,083.3	67.3	3.5 D	48.4							
3.5	19.5	4.4	24.5	28.4	48.0	1,175.4	52.7	2.5 P	55.9							
2.6	13.9	4.3	22.7	28.6	51.6	1,123.4	48.0	16 P	62.0							
2.4	11.2	4.2	19.9	29.0	60.3	1,011.8	34.1	16 P	76.7							
2.3	9.6	4.1	17.1	28.8	57.2	908.1	21.4	10 P	77.3							
2.3	8.0	4.1	15.8	28.7	56.1	1,039.4	37.9	5 P	71.1							
2.4	7.7	4.1	15.3	28.7	56.7	967.8	31.1	4 P	64.6							
2.5	8.0	4.3	18.4	28.7	57.5	938.7	23.8	10.5 P	63.6							
3.6	16.4	4.5	22.0	28.4	53.5	1,013.9	35.4	11.5 P	72.8							
3.4	13.6	4.5	25.0	27.9	43.0	991.5	33.1	16.5 D	73.6							
2.9	9.6	4.6	26.9	28.4	56.3	1,034.6	35.6	7.5 D	40.3							
2.3	5.3	4.6	31.1	28.7	63.3	970.2	26.6	8 D	50.6							
2.4	5.6	4.6	33.5	28.7	65.4	924.6	21.1	10.5 D	52.6							
2.5	6.0	4.6	35.2	28.3	60.8	962.7	27.9	11 D	50.0							
2.5	6.3	4.8	40.5	28.0	54.3	910.6	20.8	17.5 D	48.7							
2.6	7.4	4.9	43.7	27.8	49.3	948.0	25.9	19 D	41.9							
3.7	13.6	5.3	49.5	27.7	47.7	931.1	23.9	34.5 D	40.1							
3.0	12.3	5.3	51.8	27.6	42.6	956.8	29.0	37.5 D	32.7	$249						
4.1	20.7	5.3	55.4	27.2	32.8	880.7	19.2	36.5 D	18.8	1,477						
4.2	23.4	5.1	57.5	27.0	28.8	1,033.6	38.6	25 D	19.1	2,620						
4.3	30.6	5.3	64.7	27.1	31.9	1,058.7	44.3	26 D	34.7	2,589						
4.2	29.6	5.3	63.2	27.5	37.4	1,066.1	36.9	33 D	33.5	3,434						
4.5	27.9	[c]5.2	[c]61.7	27.3	33.0	1,133.9	59.0	32 D	26.1	3,489						
4.0	24.4	[c]5.1	[c]61.5	27.3	33.0	1,094.1	46.4	29.5 D	27.3	3,883						
3.6	19.4	[c]4.9	[c]55.2	27.5	34.1	1,132.3	49.6	27.5 D	29.0	2,543						
6.5	29.3	[c]4.9	[c]51.4	27.6	36.4	1,144.0	50.1	31 D	30.8	3,014						
7.1	32.9	[c]4.9	[c]48.9	27.2	27.5	1,140.7	54.2	29 D	24.2	3,685						
5.4	30.3	[c]4.9	[c]51.3	27.1	22.7	1,077.6	45.3	20 D	27.6	2,700						
4.8	26.1	[c]5.0	[c]53.5	27.4	29.4	1,283.9	65.7	4.5 D	36.9	2,686						
4.2	26.1	[c]4.7	[c]46.0	27.8	36.1	1,177.0	55.6	13 P	53.4	1,530						
4.0	26.8	4.8	48.6	27.6	32.3	1,107.7	48.1	2.5 D	53.2	213						
4.9	30.3	[c]4.7	[c]47.8	27.2	24.9	1,191.3	65.2	11.5 D	47.3	836						
5.5	39.2	[c]4.8	[c]51.6	27.4	29.4	1,222.4	63.5	5 P	64.7		515	52.4				
6.6	46.1	[c]4.8	[c]49.3	27.5	32.8	1,202.1	60.8	3.5 P	65.3		60	47.8				
7.4	49.3	[c]4.9	[c]52.9	27.7	33.3	1,015.5	38.8	3.5 P	65.1		2,188	61.7				

AER 1911

Consensus général croissant dans le milieu académique

- **Principes FAIR**
- Principes de citation de données
- Reproductibilité computationnelle

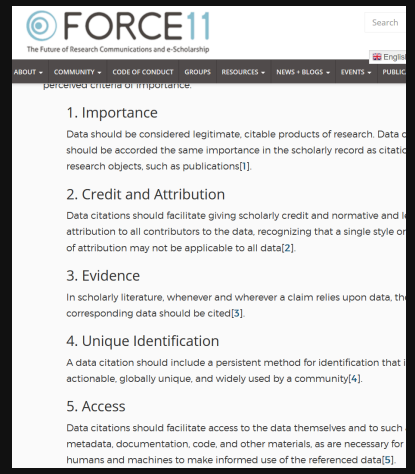
- **F**indable (trouvable)
- **A**ccessible
- **I**nteroperable (interopérable)
- **R**eusable (réutilisable)



Principes de citation de données

- Principes FAIR
- **Principes de citation de données**
- Reproductibilité computationnelle

- Pour les rendre **trouvables**, ***citations***,
- Donner **attribution** et ***crédit*** pour les données.



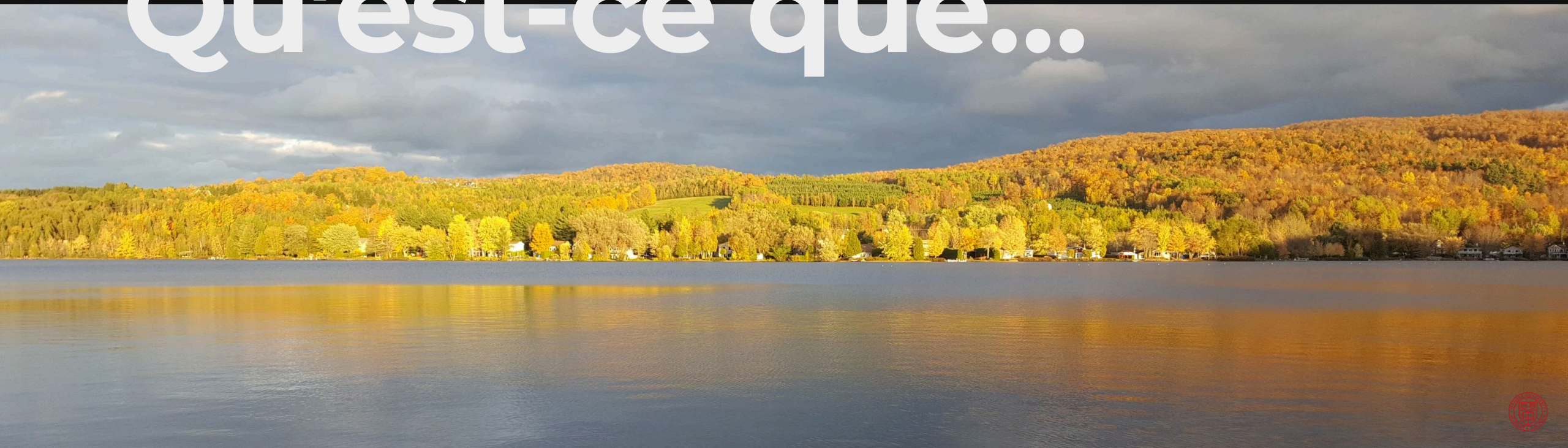
Reproductibilité computationnelle

- Principes FAIR
- Principes de citation de données
- **Reproductibilité computationnelle**

- Sujet principal d'aujourd'hui

> ***Reproductibilité*** signifie obtenir des résultats computationnels cohérents en utilisant les mêmes données d'entrée, étapes computationnelles, méthodes, code et conditions d'analyse.[2]

Qu’est-ce que…



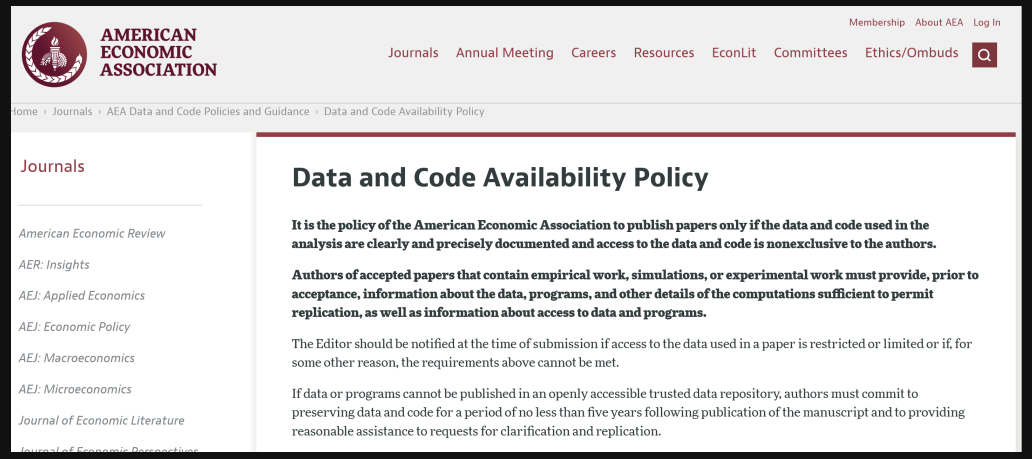
Qu’est-ce qu’un paquet de réplication ?

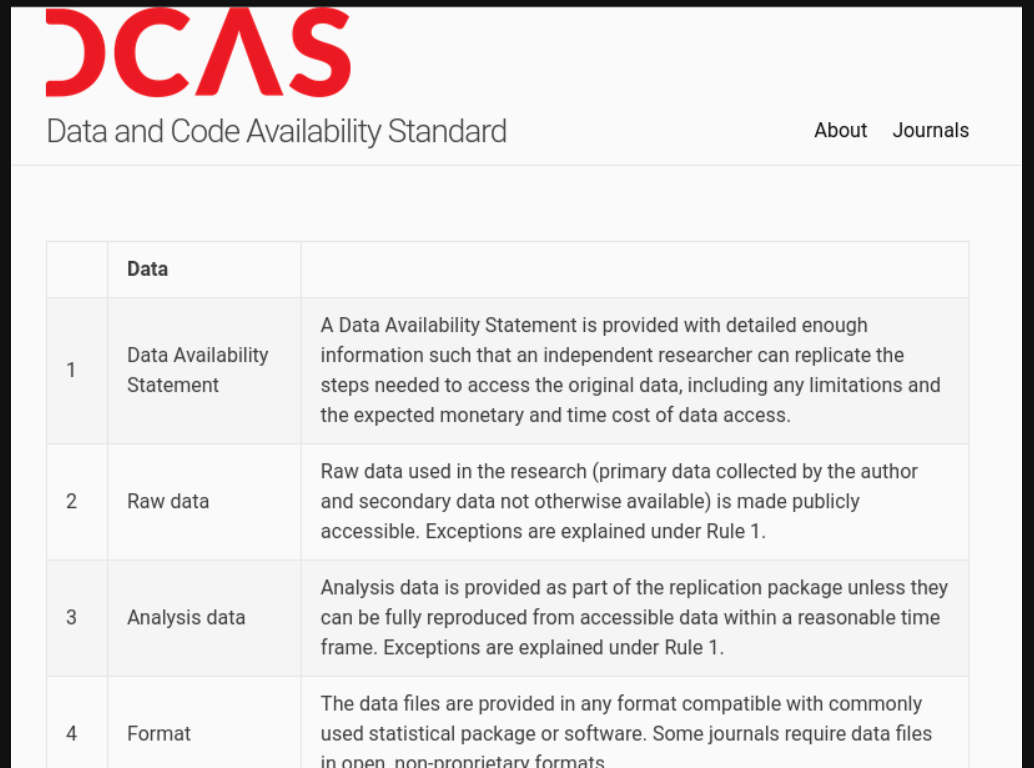
Un paquet de réplication comprend

- Code
- Données
- Matériaux (pour les sondages, expériences, ...)
- Instructions sur la façon d'obtenir les données non incluses
- Instructions sur la façon de tout combiner
- Problèmes connus documentés

Conforme à...

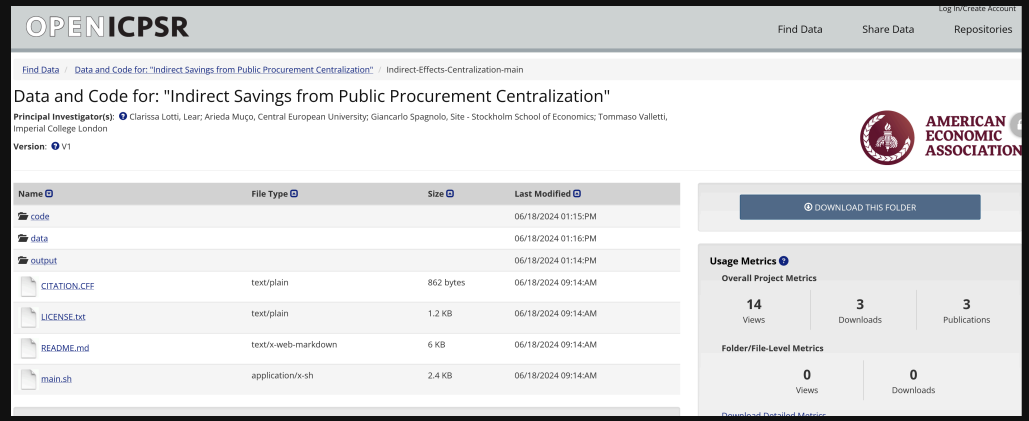
- Politique de disponibilité des données et du code de l'AEA
- Norme de disponibilité des données et du code DCAS v1.0 .





Est stocké dans…

- Dépôt de données et de code de l’AEA
- Autres dépôts de confiance



Meilleures pratiques ?

En résumé

- **Pourquoi**
  - Crédibilité
  - Transparence (ouverture)
  - Efficacité du discours scientifique ([exemple])
- ***Comment***
  - Principes FAIR
  - Principes de citation de données
  - Reproductibilité computationnelle

- Sous forme de **paquets de réplication**
  - Code
  - Données
  - Matériaux (pour les sondages, expériences, …)
  - Instructions sur la façon d’obtenir les données non incluses
  - Instructions sur la façon de tout combiner
  - Problèmes connus documentés

Qui ?

Qui ?

- 🐇 Auteurs à l'**acceptation conditionnelle**
- 🐢 Auteurs à la ***soumission***
- 🐁 Auteurs au **début** du projet

- 👴🏽👵🏽 Chercheurs expérimentés
- 👶🏽👶🏻 Chercheurs juniors
- 👨‍🎓👩‍🎓 Doctorants
- 🧒👦 Étudiants de premier cycle

Qui ?

Vous.

Vous

👶 Maintenant :

- développement plus efficace
- collaboration plus efficace
- plus d’assurance que “tout fonctionne”

👵🏽 Bientôt

- développement plus efficace entre les projets
- réponse plus efficace aux éditeurs et réviseurs
- … pendant que vous êtes dans une **nouvelle** institution, sur un ***nouvel ordinateur***, avec trois cours à **préparer**, et (luxe

Vous

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Experience of irreproducibility as a risk factor for poor mental health in biomedical science doctoral students: A survey and interview-based study

Nasser Lubega, Abigail Anderson, Nicole C. Nelson

Published: November 15, 2023 • https://doi.org/10.1371/journal.pone.0293584


Comment ?

Comment créer une recherche reproductible ?

Habitudes

- Reproductibilité **dès le jour 1**
- Adopter des ***habitudes*** reproductibles
- **Prendre des notes** pendant que vous faites les choses, pas après
- Utiliser le ***contrôle de version***

Stratégie

Empathie computationnelle : pensez à la prochaine personne qui exécutera ceci - Cela pourrait être **vous** dans 5 ans !

Hands-on: A very imperfect example

Presentation of the example

Day 1 reproducibility

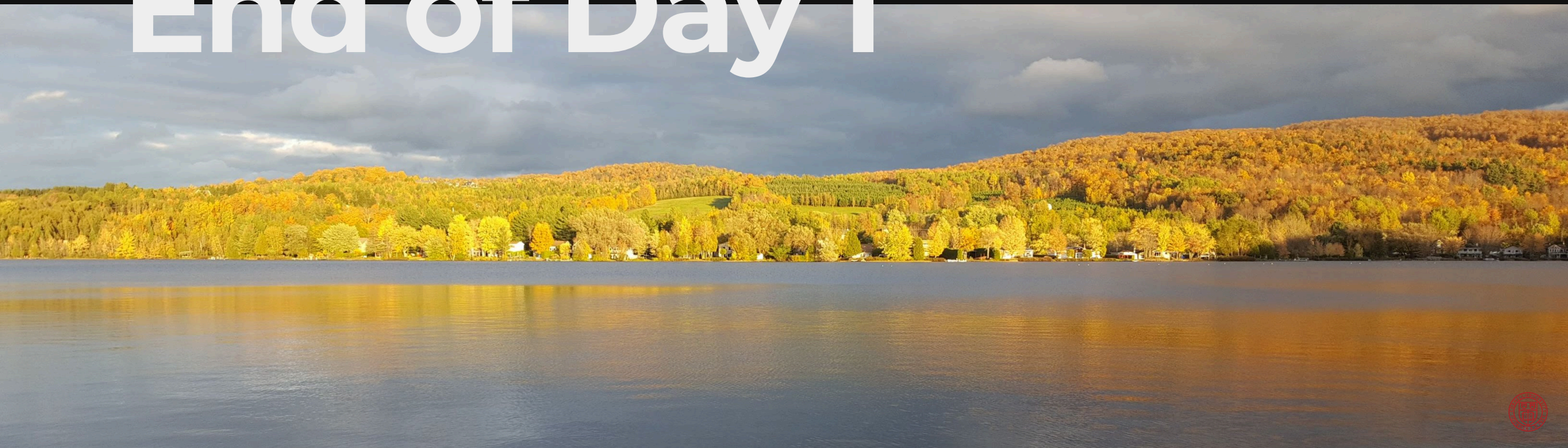
Day 1: How to not to organize your work

A tale of ineffective technical collaboration

Day 1: Setting yourself up for reproducibility

An approach to be reproducible from Day 1

End of Day 1



Appendix

Où aller ?

Choix

- Question No. 1 : Accès aux données et préservation 💥 [a]
- Question No. 2 : Données confidentielles - même chose ⬆️ ! [a] [b]
- Vérification du cycle de vie : Auto-vérification de la reproductibilité et présentation
- Nouveaux défis : IA et mégadonnées mais même chose ⬆️ !
- Nouvelles méthodes : Transparence externalisée ou certifiée
- Mise en œuvre dans le milieu académique : Étudiants !

Resources

README

> Lars Vilhuber, Connolly, M., Koren, M., Llull, J., & Morrow, P. (2022). A template README for social science replication packages (v1.1). Social Science Data Editors. https://doi.org/10.5281/zenodo.7293838

You can download the Word, LaTeX, or Markdown version of the README with lots of examples.

Other guidance

- Presentation on “Self Checking Reproducibility” and its associated website
- Guidance when (some) data are confidential: https://labordynamicsinstitute.github.io/reproducibility-confidential/
- Guidance for citations: https://social-science-data-editors.github.io/guidance/addtl-data-citation-guidance.html

Extra info

- This document's source: https://github.com/larsvilhuber/ecole-d-ete-ciqss-2026
- Licensed under  CC BY-NC 4.0

Sources

- Images: NYT, Bluesky 1, 2, Ike Hayman/Wikimedia

Détails sur la transparence, etc.

Transparence

- Provenance des ***données***
- Traitement des données, des données brutes aux résultats (***code***)

> Prior to acceptance, authors of papers [...] must provide the data, code, and other details of the computations **sufficient to permit replication**. These materials must be made **available and retained in an openly accessible** trusted data repository, such as the AEA Data and Code Repository.

Transparence

- Provenance des ***données***
- Traitement des données, des données brutes aux résultats (***code***)

> DCAS v1.0 The Canadian Journal of Economics endorses DCAS, the Data and Code Availability Standard [v1.0], and its data and code availability policy is compatible with DCAS.

Exhaustivité

- Toutes les données doivent être identifiées et l’accès décrit
- Tout le code doit être décrit et fourni
- Tous les matériaux doivent être fournis (formulaires de sondage, etc.)

> Les auteurs … doivent fournir, avant l’acceptation, les **données, programmes et autres détails** des calculs **suffisants** pour permettre la réplication

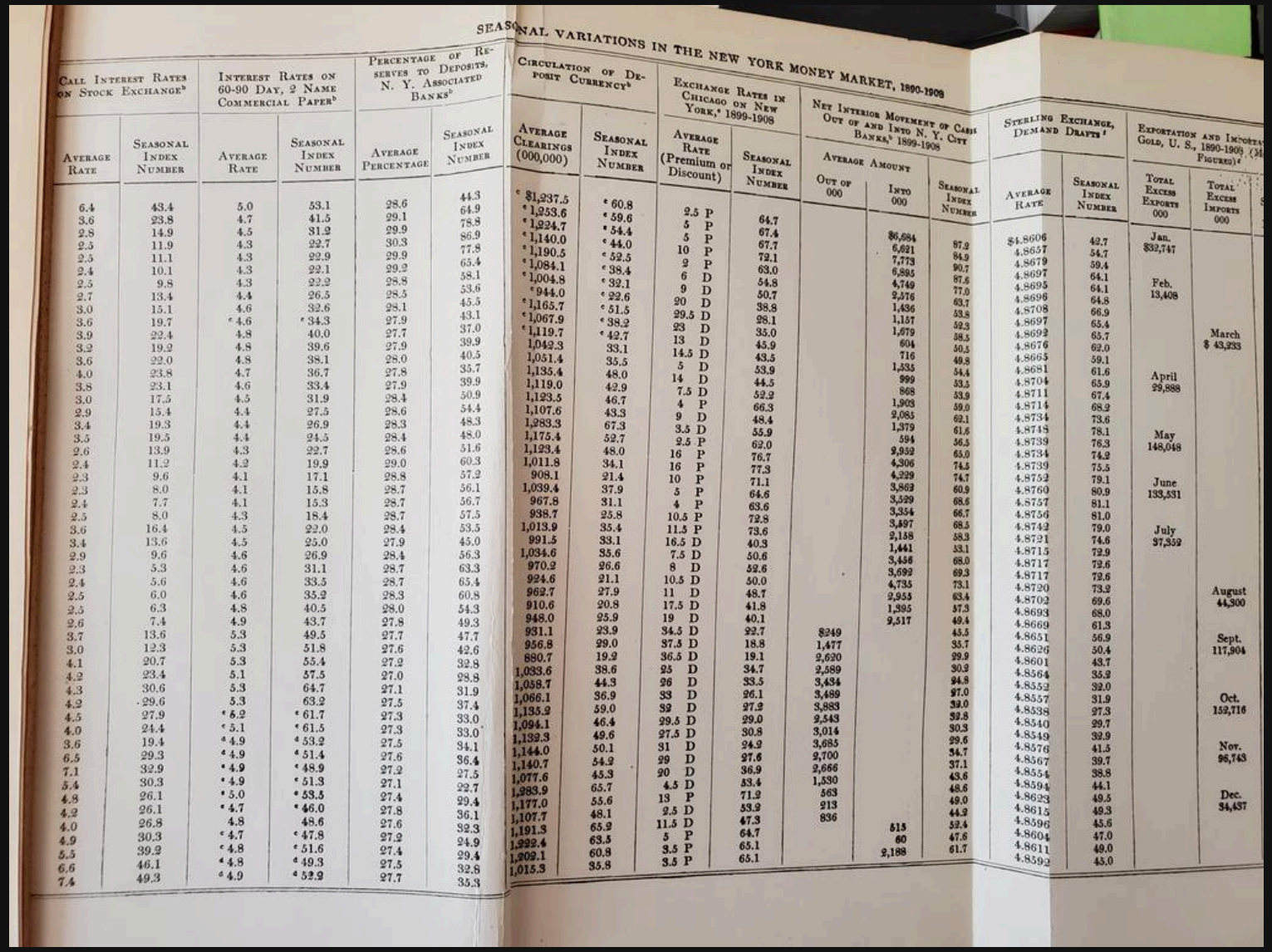
Préservation

- Toutes les ***données*** doivent être préservées pour les futurs réplicateurs
  - Idéalement, dans le paquet de réplication, sous réserve des conditions d’utilisation, pour plus de commodité
  - Sinon, dans un **dépôt de confiance**

Préservation

- Le ***code*** doit être dans un dépôt de confiance
  - Habituellement, dans le paquet de réplication
  - Les sites Web, Github, ne sont ***pas acceptables***

Historiquement

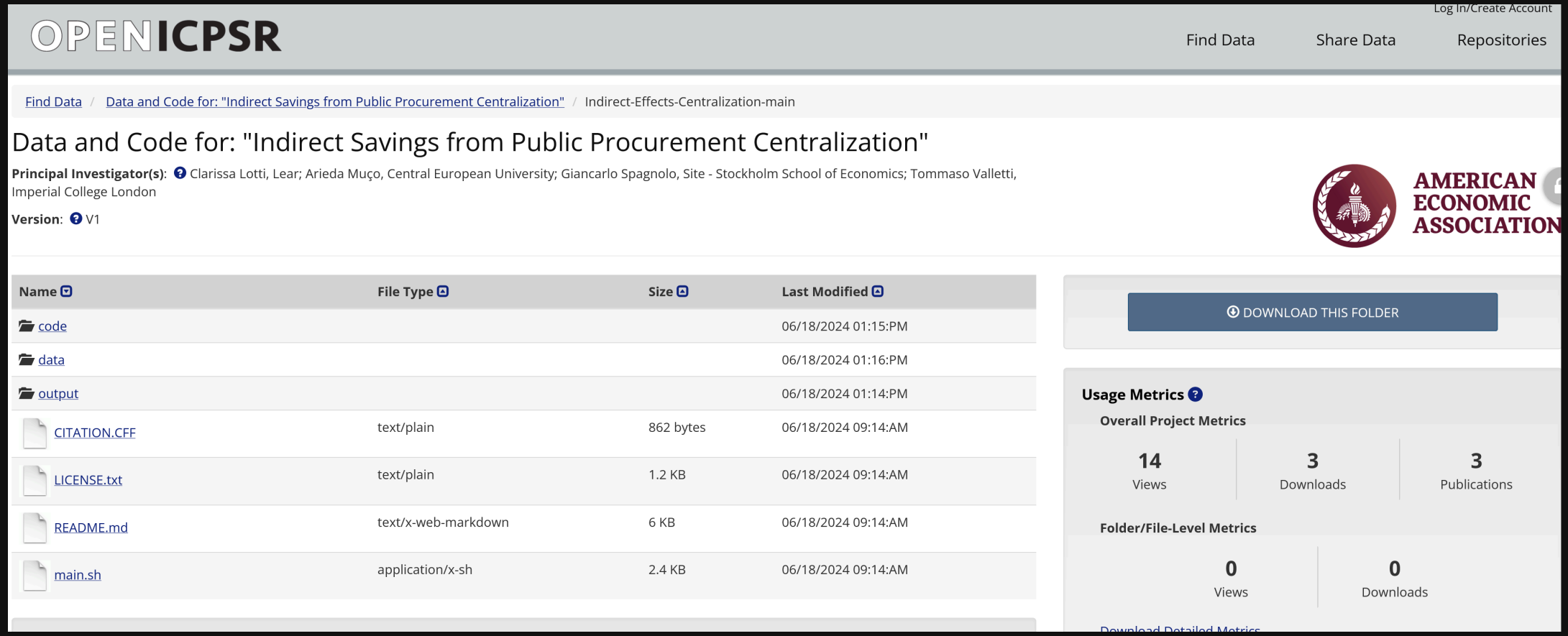


SEASONAL VARIATIONS IN THE NEW YORK MONEY MARKET, 1890-1908

Call Interest Rates on Stock Exchange[b]		Interest Rates on 60-90 Day, 2 Name Commercial Paper[b]		Percentage of Reserves to Deposits, N. Y. Associated Banks[b]		Circulation of Deposit Currency[b]		Exchange Rates in Chicago on New York,[c] 1899-1908		Net Interior Movement of Cash Out of and Into N. Y. City Banks,[b] 1899-1908			Sterling Exchange, Demand Drafts[d]		Exportation and Importation Gold, U. S., 1890-1908 (Minus Figures)[e]	
Average Rate	Seasonal Index Number	Average Rate	Seasonal Index Number	Average Percentage	Seasonal Index Number	Average Clearings (000,000)	Seasonal Index Number	Average Rate (Premium or Discount)	Seasonal Index Number	Average Amount Out of 000	Average Amount Into 000	Seasonal Index Number	Average Rate	Seasonal Index Number	Total Excess Exports 000	Total Excess Imports 000
6.4	43.4	5.0	53.1	28.6	44.3	[c] $1,237.5	[c] 60.8	2.5 P	64.7				$4.8606	42.7	Jan.	
3.6	23.8	4.7	41.5	29.1	64.9	[c] 1,253.6	[c] 59.6	5 P	67.4		86,684	87.2	4.8657	54.7	$32,747	
2.8	14.9	4.5	31.2	29.9	78.8	[c] 1,224.7	[c] 54.4	5 P	67.7		6,621	84.9	4.8679	59.4		
2.5	11.9	4.3	22.7	30.3	86.9	[c] 1,140.0	[c] 44.0	10 P	72.1		7,773	90.7	4.8697	64.1	Feb.	
2.5	11.1	4.3	22.9	29.9	77.8	[c] 1,190.5	[c] 52.5	2 P	63.0		6,895	87.6	4.8695	64.1	13,408	
2.4	10.1	4.3	22.1	29.2	65.4	[c] 1,084.1	[c] 38.4	6 D	54.8		4,749	77.0	4.8696	64.8		
2.5	9.8	4.3	22.2	28.8	58.1	[c] 1,004.8	[c] 32.1	9 D	50.7		2,576	63.7	4.8708	66.9		
2.7	13.4	4.4	26.5	28.5	53.6	[c] 944.0	[c] 22.6	20 D	38.8		1,436	53.8	4.8697	65.4		
3.0	15.1	4.6	32.6	28.1	45.5	[c] 1,165.7	[c] 51.5	29.5 D	28.1		1,157	52.3	4.8692	65.7		March
3.6	19.7	[c] 4.6	[c] 34.3	27.9	43.1	[c] 1,067.9	[c] 38.2	23 D	35.0		1,679	58.5	4.8676	62.0		$ 43,233
3.9	22.4	4.8	40.0	27.7	37.0	[c] 1,119.7	[c] 42.7	13 D	45.9		604	50.5	4.8665	59.1		
3.2	19.2	4.8	39.6	27.9	39.9	1,042.3	33.1	14.5 D	43.5		716	49.8	4.8681	61.6		
3.6	22.0	4.8	38.1	28.0	40.5	1,051.4	35.5	5 D	53.9		1,535	54.4	4.8704	65.9	April	
4.0	23.8	4.7	36.7	27.8	35.7	1,135.4	48.0	14 D	44.5		999	53.5	4.8711	67.4	29,888	
3.8	23.1	4.6	33.4	27.9	39.9	1,119.0	42.9	7.5 D	52.2		868	53.9	4.8714	68.2		
3.0	17.5	4.5	31.9	28.4	50.9	1,123.5	46.7	4 P	66.3		1,903	59.0	4.8734	73.6		
2.9	15.4	4.4	27.5	28.6	54.4	1,107.6	43.3	9 D	48.4		2,085	62.1	4.8748	78.1		
3.4	19.3	4.4	26.9	28.3	48.3	1,283.3	67.3	3.5 D	55.9		1,379	61.6	4.8739	76.3	May	
3.5	19.5	4.4	24.5	28.4	48.0	1,175.4	52.7	2.5 P	62.0		594	56.5	4.8734	74.2	148,048	
2.6	13.9	4.3	22.7	28.6	51.6	1,123.4	48.0	16 P	76.7		2,952	65.0	4.8739	75.5		
2.4	11.2	4.2	19.9	29.0	60.3	1,011.8	34.1	16 P	77.3		4,306	74.5	4.8752	79.1	June	
2.3	9.6	4.1	17.1	28.8	57.2	908.1	21.4	10 P	71.1		4,229	74.7	4.8760	80.9	133,331	
2.3	8.0	4.1	15.8	28.7	56.1	1,039.4	37.9	5 P	64.6		3,863	60.9	4.8757	81.1		
2.4	7.7	4.1	15.3	28.7	56.7	967.8	31.1	4 P	63.6		3,529	68.6	4.8756	81.0		
2.5	8.0	4.3	18.4	28.7	57.5	938.7	25.8	10.5 P	72.8		3,354	66.7	4.8742	79.0	July	
3.6	16.4	4.5	22.0	28.4	53.5	1,013.9	35.4	11.5 P	73.6		3,597	68.5	4.8721	74.6	37,352	
3.4	13.6	4.5	25.0	27.9	45.0	991.5	33.1	16.5 D	40.3		2,158	58.3	4.8715	72.9		
2.9	9.6	4.6	26.9	28.4	56.3	1,034.6	35.6	7.5 D	50.6		1,441	53.1	4.8717	72.6		
2.3	5.3	4.6	31.1	28.7	63.3	970.2	26.6	8 D	52.6		3,456	68.0	4.8717	72.6		
2.4	5.6	4.6	33.5	28.7	65.4	924.6	21.1	10.5 D	50.0		3,592	69.3	4.8720	73.2		August
2.5	6.0	4.6	35.2	28.3	60.8	962.7	27.9	11 D	48.7		4,735	73.1	4.8702	69.6		44,300
2.5	6.3	4.8	40.5	28.0	54.3	910.6	20.8	17.5 D	41.8		2,955	63.4	4.8693	68.0		
2.6	7.4	4.9	43.7	27.8	49.3	948.0	25.9	19 D	40.1		1,595	57.3	4.8669	61.3		
3.7	13.6	5.3	49.5	27.7	47.7	931.1	23.9	34.5 D	22.7	$249	9,517	49.4	4.8651	56.9		Sept.
3.0	12.3	5.3	51.8	27.6	42.6	956.8	29.0	37.5 D	18.8	1,477		43.5	4.8626	50.4		117,904
4.1	20.7	5.3	55.4	27.2	32.8	880.7	19.2	36.5 D	19.1	2,630		33.7	4.8601	43.7		
4.2	23.4	5.1	57.5	27.0	28.8	1,033.6	38.6	25 D	34.7	2,589		29.9	4.8564	35.2		
4.3	30.6	5.3	64.7	27.1	31.9	1,058.7	44.3	26 D	33.5	3,434		30.2	4.8552	32.0		
4.2	29.6	5.3	63.2	27.5	37.4	1,066.1	36.9	33 D	26.1	3,489		24.8	4.8557	31.9		Oct.
4.5	27.9	[c] 5.2	[c] 61.7	27.3	33.0	1,135.2	59.0	32 D	27.2	3,883		27.0	4.8538	27.3		152,716
4.0	24.4	[c] 5.1	[c] 61.5	27.3	33.0	1,094.1	46.4	29.5 D	29.0	2,543		32.8	4.8540	29.7		
3.6	19.4	[c] 4.9	[c] 53.2	27.5	34.1	1,132.3	49.6	27.5 D	30.8	3,014		30.3	4.8549	32.9		
6.5	29.3	[c] 4.9	[c] 51.4	27.6	36.4	1,144.0	50.1	31 D	24.2	3,685		29.6	4.8576	41.5		Nov.
7.1	32.9	[c] 4.9	[c] 48.9	27.2	27.5	1,140.7	54.2	29 D	27.6	2,700		34.7	4.8567	39.7		96,743
5.4	30.3	[c] 4.9	[c] 51.3	27.1	22.7	1,077.6	45.3	20 D	36.9	2,666		37.1	4.8554	38.8		
4.8	26.1	[c] 5.0	[c] 53.5	27.4	29.4	1,283.9	65.7	4.5 D	53.4	1,530		43.6	4.8594	44.1		
4.2	26.1	[c] 4.7	[c] 46.0	27.8	36.1	1,177.0	55.6	13 P	71.2	563		48.6	4.8623	49.5		Dec.
4.0	26.8	4.8	48.6	27.6	32.3	1,107.7	48.1	2.5 D	59.2	213		49.0	4.8615	49.3		34,437
4.9	30.3	[c] 4.7	[c] 47.8	27.2	24.9	1,191.3	65.2	11.5 D	47.3	836		44.2	4.8596	45.6		
5.5	39.2	[c] 4.8	[c] 51.6	27.4	29.4	1,222.4	63.5	5 P	64.7		515	52.4	4.8604	47.0		
6.6	46.1	[c] 4.8	[c] 49.3	27.5	32.8	1,202.1	60.8	3.5 P	65.1		60	47.6	4.8611	49.0		
7.4	49.3	[c] 4.9	[c] 52.9	27.7	35.3	1,015.3	35.8	3.5 P	65.1		2,188	61.7	4.8592	45.0		

AER 1911 merci à Stefano Dellavigna

Préservation moderne



OPENICPSR

Log In/Create Account

Find Data Share Data Repositories

Find Data / Data and Code for: "Indirect Savings from Public Procurement Centralization" / Indirect-Effects-Centralization-main

Data and Code for: "Indirect Savings from Public Procurement Centralization"

Principal Investigator(s): Clarissa Lotti, Lear; Arieda Muço, Central European University; Giancarlo Spagnolo, Site - Stockholm School of Economics; Tommaso Valletti, Imperial College London

Version: V1

AMERICAN ECONOMIC ASSOCIATION

Name	File Type	Size	Last Modified
code			06/18/2024 01:15:PM
data			06/18/2024 01:16:PM
output			06/18/2024 01:14:PM
CITATION.CFF	text/plain	862 bytes	06/18/2024 09:14:AM
LICENSE.txt	text/plain	1.2 KB	06/18/2024 09:14:AM
README.md	text/x-web-markdown	6 KB	06/18/2024 09:14:AM
main.sh	application/x-sh	2.4 KB	06/18/2024 09:14:AM

DOWNLOAD THIS FOLDER

Usage Metrics

Overall Project Metrics

14	3	3
Views	Downloads	Publications

Folder/File-Level Metrics

0	0
Views	Downloads

Download Detailed Metrics

Données confidentielles?

- Lorsque les données n’appartiennent pas au chercheur, celui-ci n’a aucun contrôle sur la préservation, l’accès !
- Parfois, les ***conditions d’utilisation*** empêchent le chercheur de révéler les métadonnées (nom de l’entreprise, emplacement)

Transparence encore

- Cependant :
  - Pas d’exception pour le besoin de **décrire** l’accès (propre et autre)
  - Pas d’exception pour le besoin de **décrire** pleinement le traitement (éventuellement avec du code caviardé)

Reproductibilité en économie et au-delà

Social Science Data Editors

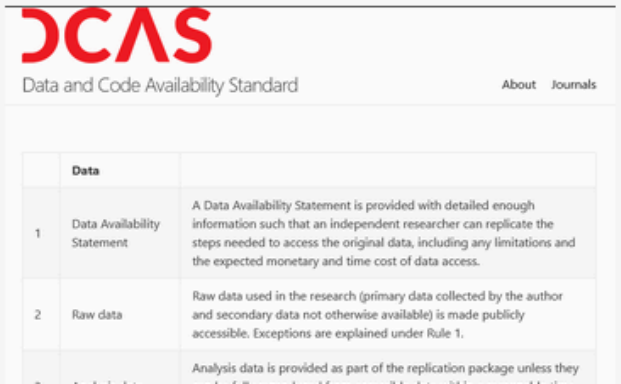
Social Science Data Editors

Improving reproducibility in the social and economic sciences

Data and Code Availability Standard

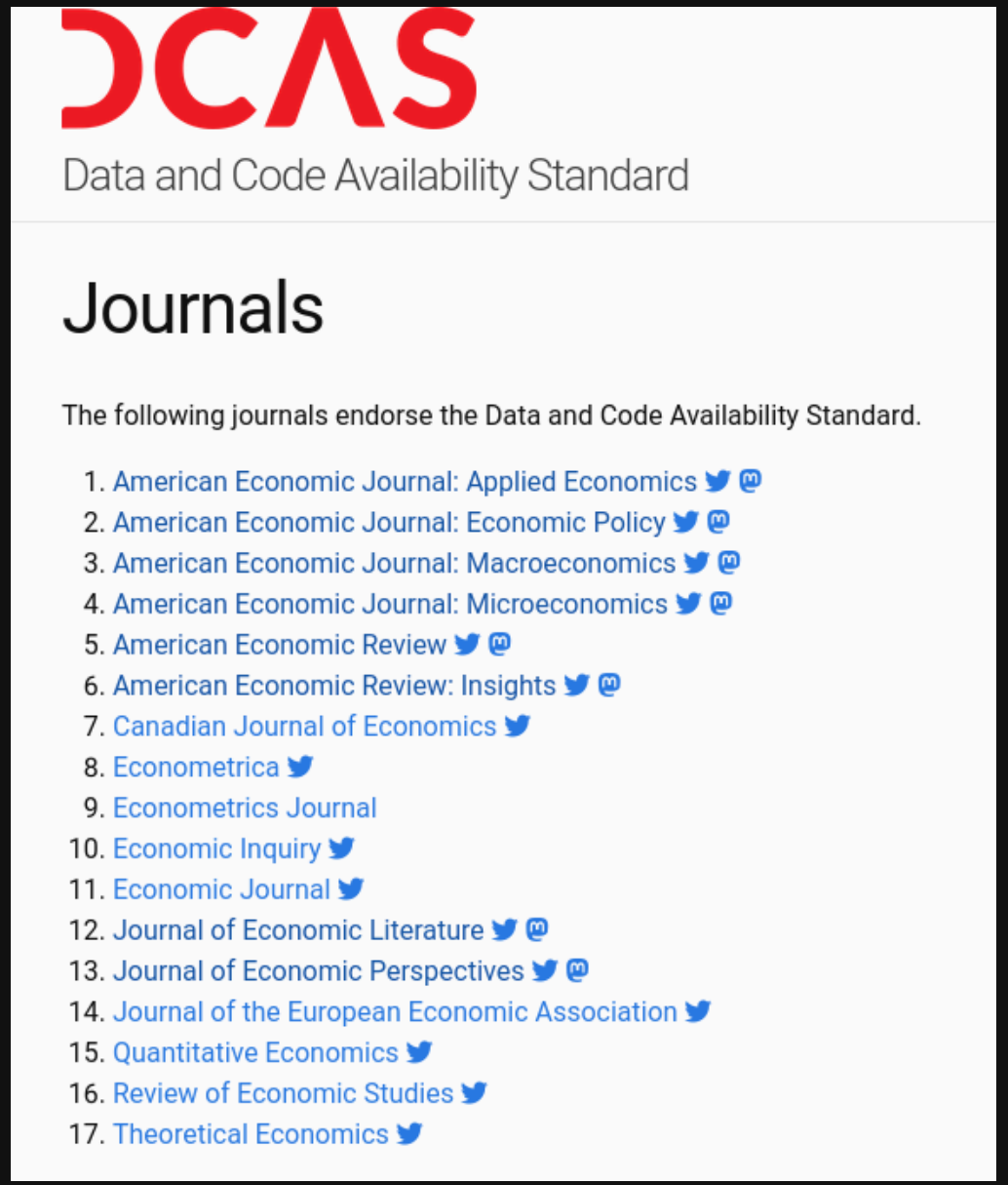
DCAS The Data and Code Availability Standard (DCAS) is a standard for sharing research code and data, endorsed by leading journals in social sciences. See https://datacodestandard.org/ for more information.

DOI 10.5281/zenodo.7436134



Éditeurs de données

- American Economic Association (8)
- Econometric Society (3)
- Canadian Journal of Economics (1)
- Royal Economic Society (2)
- Western Economic Association International (1)
- European Economic Association (1)
- Review of Economic Studies (1)
- Journal of the European Economic Association (1)
- Journal of Political Economy (3)
- American Journal of Political Science (1)



DCAS

Data and Code Availability Standard

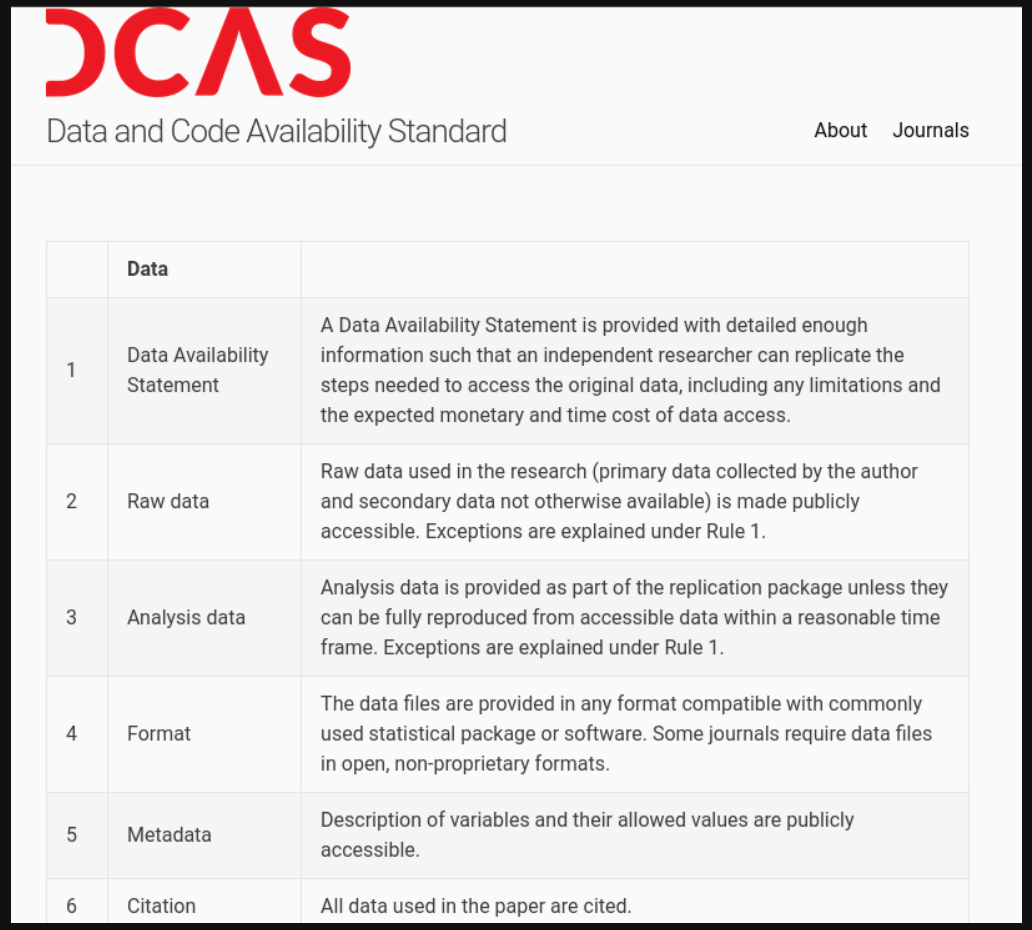
Journals

The following journals endorse the Data and Code Availability Standard.

1. American Economic Journal: Applied Economics
2. American Economic Journal: Economic Policy
3. American Economic Journal: Macroeconomics
4. American Economic Journal: Microeconomics
5. American Economic Review
6. American Economic Review: Insights
7. Canadian Journal of Economics
8. Econometrica
9. Econometrics Journal
10. Economic Inquiry
11. Economic Journal
12. Journal of Economic Literature
13. Journal of Economic Perspectives
14. Journal of the European Economic Association
15. Quantitative Economics
16. Review of Economic Studies
17. Theoretical Economics

Politiques communes

https://social-science-data-editors.github.io/

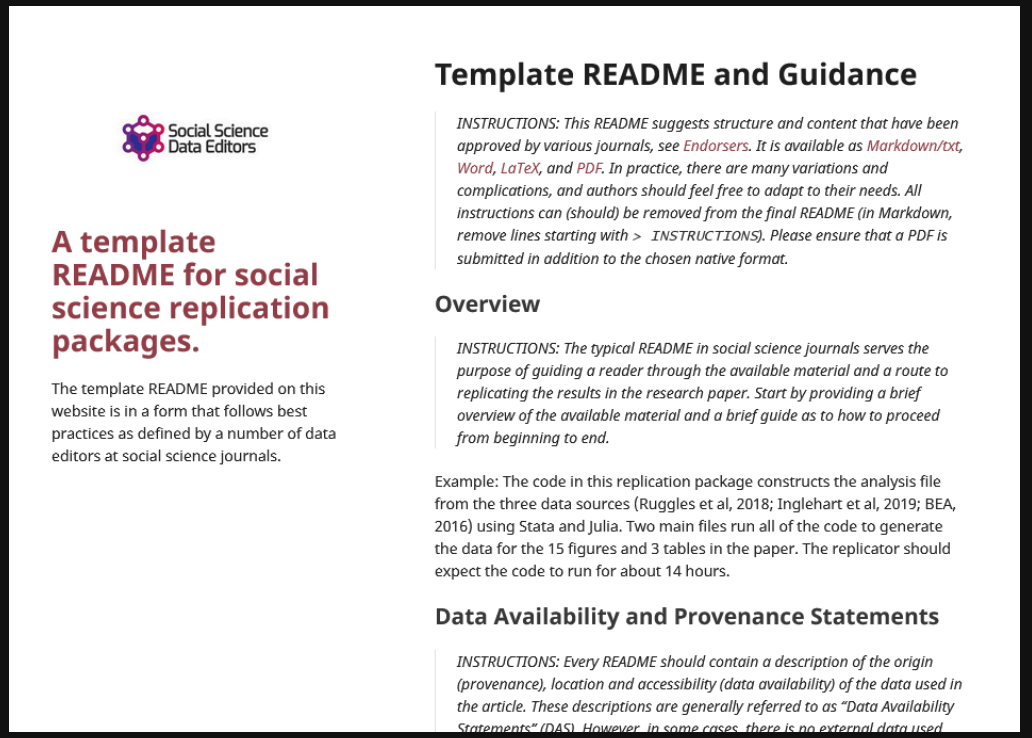


DCAS

Data and Code Availability Standard

About Journals

	Data	
1	Data Availability Statement	A Data Availability Statement is provided with detailed enough information such that an independent researcher can replicate the steps needed to access the original data, including any limitations and the expected monetary and time cost of data access.
2	Raw data	Raw data used in the research (primary data collected by the author and secondary data not otherwise available) is made publicly accessible. Exceptions are explained under Rule 1.
3	Analysis data	Analysis data is provided as part of the replication package unless they can be fully reproduced from accessible data within a reasonable time frame. Exceptions are explained under Rule 1.
4	Format	The data files are provided in any format compatible with commonly used statistical package or software. Some journals require data files in open, non-proprietary formats.
5	Metadata	Description of variables and their allowed values are publicly accessible.
6	Citation	All data used in the paper are cited.



Social Science Data Editors

A template README for social science replication packages.

The template README provided on this website is in a form that follows best practices as defined by a number of data editors at social science journals.

Template README and Guidance

INSTRUCTIONS: This README suggests structure and content that have been approved by various journals, see Endorsers. It is available as Markdown/txt, Word, LaTeX, and PDF. In practice, there are many variations and complications, and authors should feel free to adapt to their needs. All instructions can (should) be removed from the final README (in Markdown, remove lines starting with > INSTRUCTIONS). Please ensure that a PDF is submitted in addition to the chosen native format.

Overview

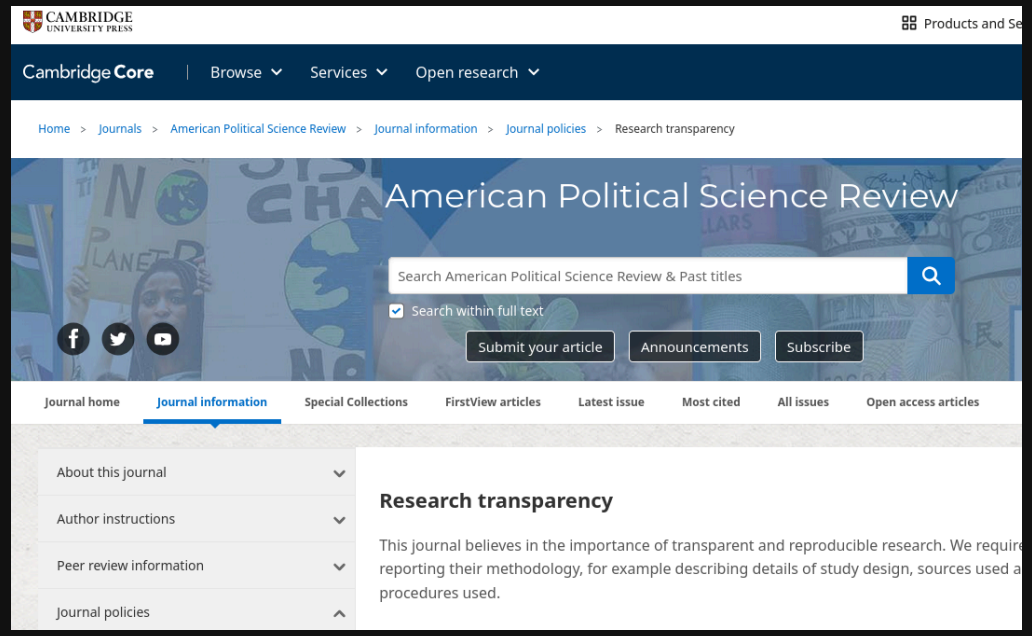
INSTRUCTIONS: The typical README in social science journals serves the purpose of guiding a reader through the available material and a route to replicating the results in the research paper. Start by providing a brief overview of the available material and a brief guide as to how to proceed from beginning to end.

Example: The code in this replication package constructs the analysis file from the three data sources (Ruggles et al, 2018; Inglehart et al, 2019; BEA, 2016) using Stata and Julia. Two main files run all of the code to generate the data for the 15 figures and 3 tables in the paper. The replicator should expect the code to run for about 14 hours.

Data Availability and Provenance Statements

INSTRUCTIONS: Every README should contain a description of the origin (provenance), location and accessibility (data availability) of the data used in the article. These descriptions are generally referred to as "Data Availability Statements" (DAS). However, in some cases, there is no external data used

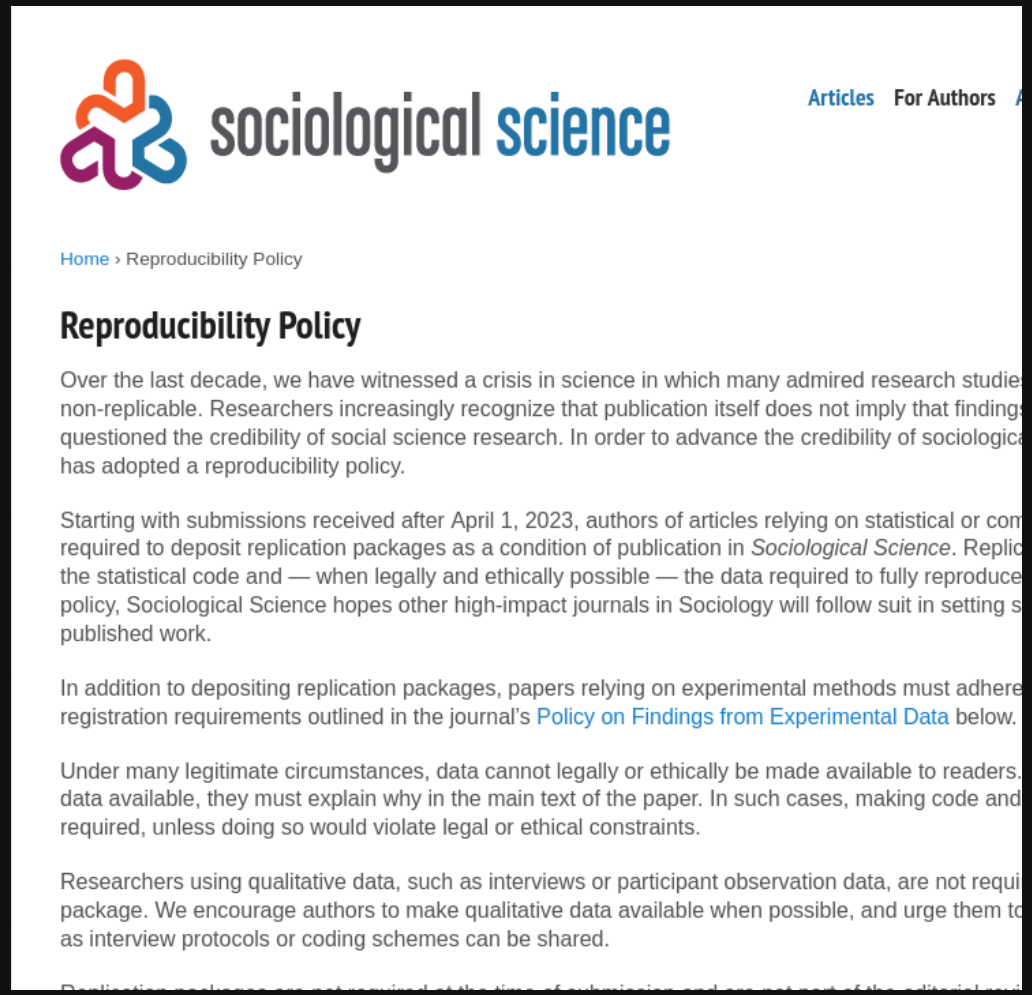
Ailleurs : Science politique



APSR

AJPS

Ailleurs : Sociologie



sociological science

Articles For Authors

Home › Reproducibility Policy

Reproducibility Policy

Over the last decade, we have witnessed a crisis in science in which many admired research studie non-replicable. Researchers increasingly recognize that publication itself does not imply that finding questioned the credibility of social science research. In order to advance the credibility of sociologica has adopted a reproducibility policy.

Starting with submissions received after April 1, 2023, authors of articles relying on statistical or com required to deposit replication packages as a condition of publication in *Sociological Science*. Replic the statistical code and — when legally and ethically possible — the data required to fully reproduce policy, Sociological Science hopes other high-impact journals in Sociology will follow suit in setting s published work.

In addition to depositing replication packages, papers relying on experimental methods must adhere registration requirements outlined in the journal's Policy on Findings from Experimental Data below.

Under many legitimate circumstances, data cannot legally or ethically be made available to readers. data available, they must explain why in the main text of the paper. In such cases, making code and required, unless doing so would violate legal or ethical constraints.

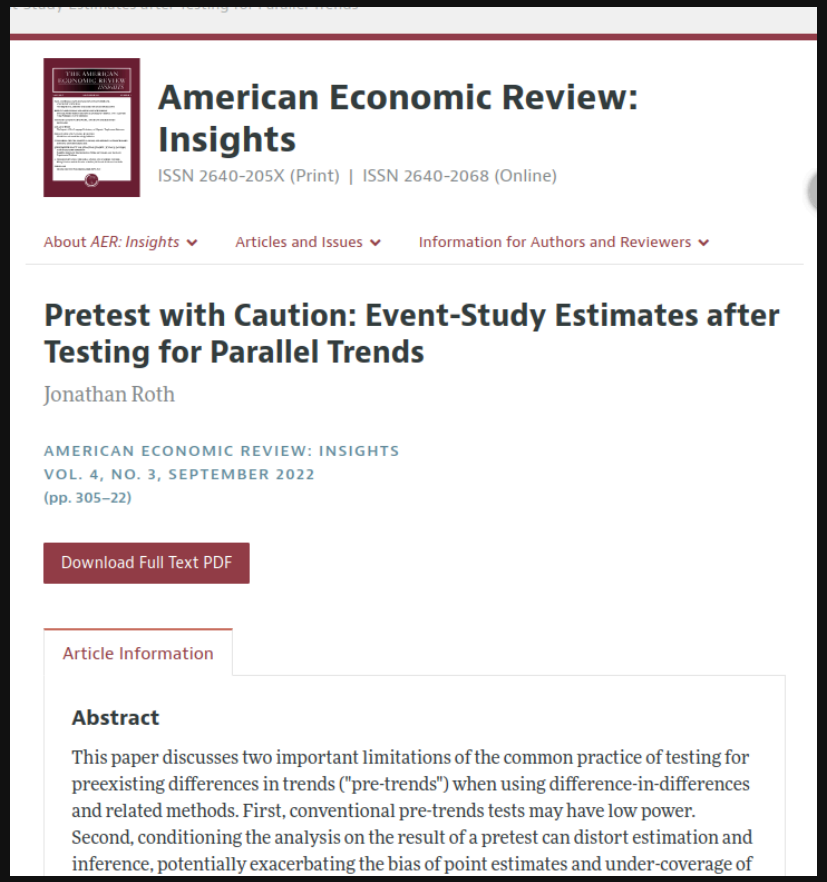
Researchers using qualitative data, such as interviews or participant observation data, are not requi package. We encourage authors to make qualitative data available when possible, and urge them to as interview protocols or coding schemes can be shared.

Sociological Science

Avantages

S’appuyer sur le travail des autres

Notes : “***J’exclus 43 articles pour lesquels les données pour reproduire le graphique principal de l’étude d’événement n’étaient pas disponibles.***”



American Economic Review: Insights

ISSN 2640-205X (Print) | ISSN 2640-2068 (Online)

About *AER: Insights* | Articles and Issues | Information for Authors and Reviewers

Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends

Jonathan Roth

AMERICAN ECONOMIC REVIEW: INSIGHTS
VOL. 4, NO. 3, SEPTEMBER 2022
(pp. 305–22)

Download Full Text PDF

Article Information

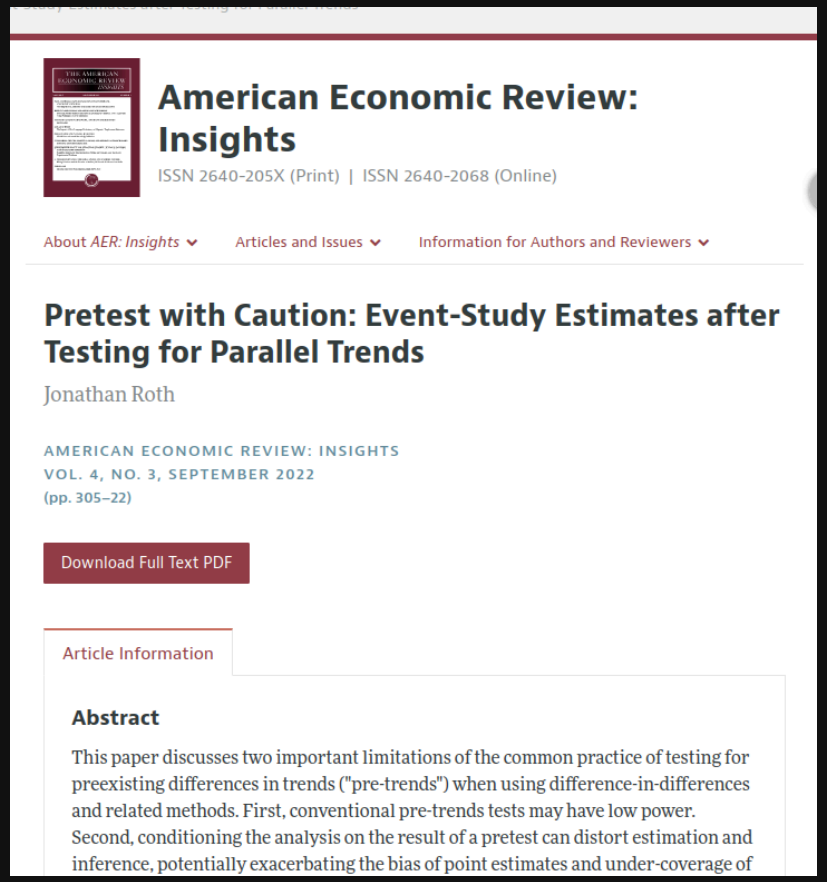
Abstract

This paper discusses two important limitations of the common practice of testing for preexisting differences in trends ("pre-trends") when using difference-in-differences and related methods. First, conventional pre-trends tests may have low power. Second, conditioning the analysis on the result of a pretest can distort estimation and inference, potentially exacerbating the bias of point estimates and under-coverage of

Roth 2022

S’appuyer sur le travail des autres

Notes : “***J’exclus 43 articles pour lesquels les données pour reproduire le graphique principal de l’étude d’événement n’étaient pas disponibles.***”



American Economic Review: Insights

ISSN 2640-205X (Print) | ISSN 2640-2068 (Online)

About *AER: Insights* | Articles and Issues | Information for Authors and Reviewers

Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends

Jonathan Roth

AMERICAN ECONOMIC REVIEW: INSIGHTS
VOL. 4, NO. 3, SEPTEMBER 2022
(pp. 305–22)

Download Full Text PDF

Article Information

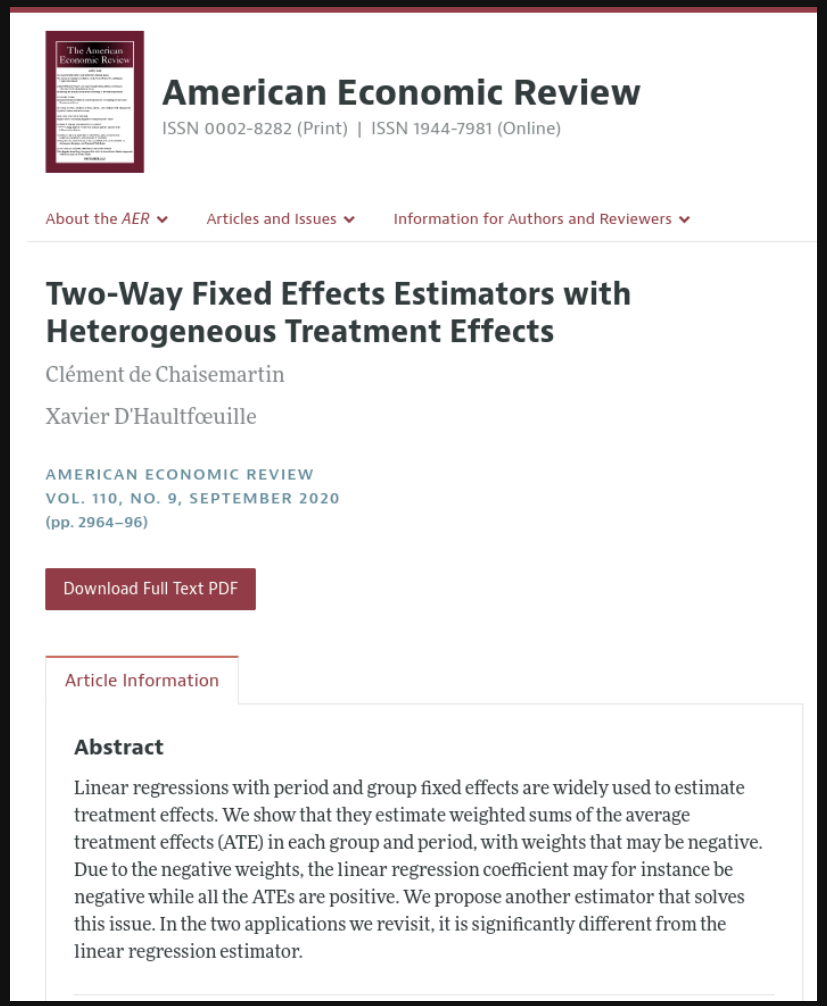
Abstract

This paper discusses two important limitations of the common practice of testing for preexisting differences in trends ("pre-trends") when using difference-in-differences and related methods. First, conventional pre-trends tests may have low power. Second, conditioning the analysis on the result of a pretest can distort estimation and inference, potentially exacerbating the bias of point estimates and under-coverage of

Roth 2022

S’appuyer sur le travail des autres : dCdH 2020

de Chaisemartin, Clément, and Xavier D’Haultfœuille. 2020. “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects.” American Economic Review 110 (9): 2964–96. DOI: 10.1257/aer.20181169



American Economic Review
ISSN 0002-8282 (Print) | ISSN 1944-7981 (Online)

About the *AER* | Articles and Issues | Information for Authors and Reviewers

Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects

Clément de Chaisemartin

Xavier D'Haultfœuille

AMERICAN ECONOMIC REVIEW
VOL. 110, NO. 9, SEPTEMBER 2020
(pp. 2964–96)

Download Full Text PDF

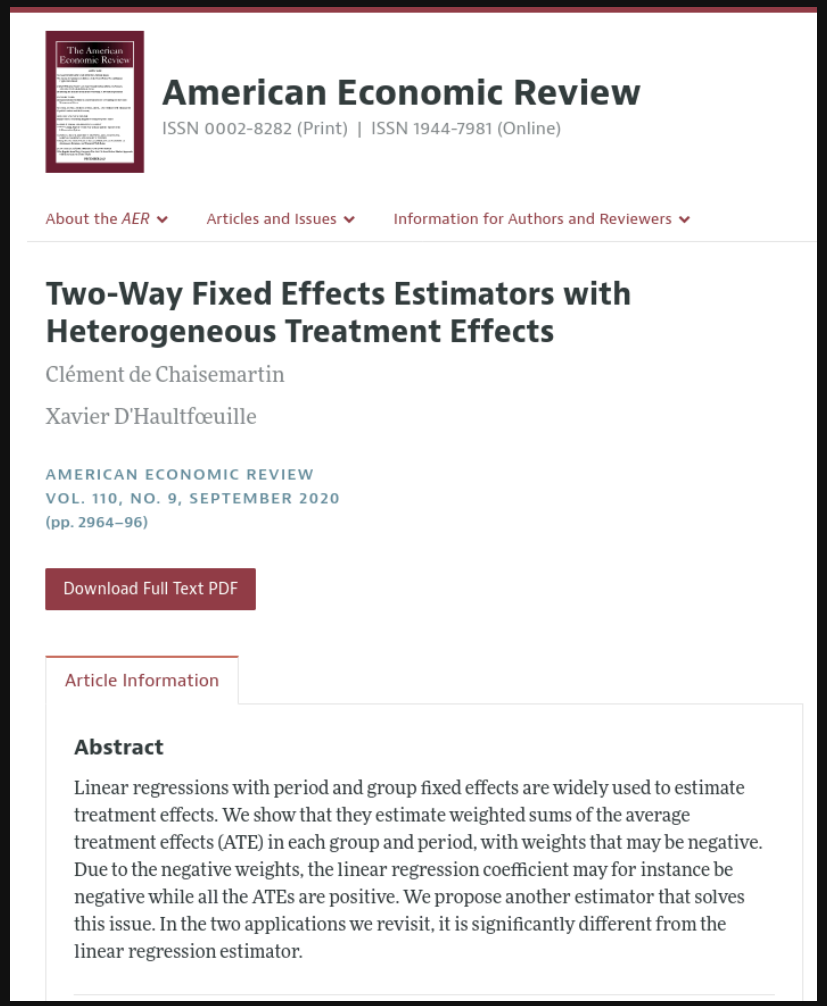
Article Information

Abstract

Linear regressions with period and group fixed effects are widely used to estimate treatment effects. We show that they estimate weighted sums of the average treatment effects (ATE) in each group and period, with weights that may be negative. Due to the negative weights, the linear regression coefficient may for instance be negative while all the ATEs are positive. We propose another estimator that solves this issue. In the two applications we revisit, it is significantly different from the linear regression estimator.

S’appuyer sur le travail des autres : dCdH 2020

Les résultats de divers autres articles sont recalculés pour démontrer empiriquement la pertinence des méthodes proposées.



American Economic Review

ISSN 0002-8282 (Print) | ISSN 1944-7981 (Online)

About the *AER* | Articles and Issues | Information for Authors and Reviewers

Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects

Clément de Chaisemartin

Xavier D'Haultfœuille

AMERICAN ECONOMIC REVIEW
VOL. 110, NO. 9, SEPTEMBER 2020
(pp. 2964–96)

Download Full Text PDF

Article Information

Abstract

Linear regressions with period and group fixed effects are widely used to estimate treatment effects. We show that they estimate weighted sums of the average treatment effects (ATE) in each group and period, with weights that may be negative. Due to the negative weights, the linear regression coefficient may for instance be negative while all the ATEs are positive. We propose another estimator that solves this issue. In the two applications we revisit, it is significantly different from the linear regression estimator.

Transparence ailleurs

Transparence externalisée

- Parlez à Limor !
- R-squared de Cornell
- cascad
- Banque mondiale

Transparence externalisée

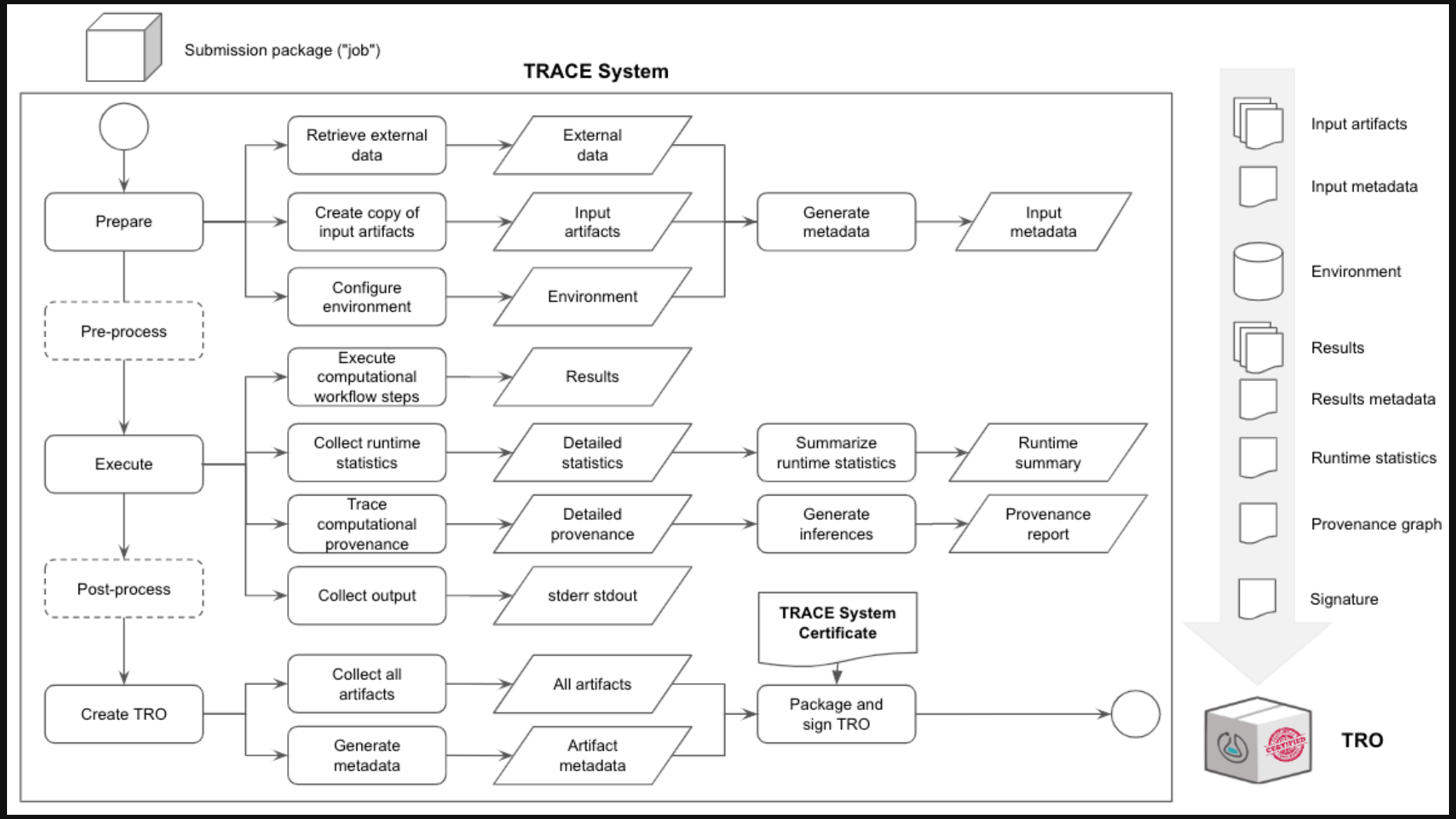
- Un tiers effectue la reproductibilité, pas vous, pas moi.
- Besoin d’une compréhension commune, de protocoles, etc.
- Protocole de l’AEA
- Nous faisons cela environ une douzaine de fois par an

Transparence externalisée

Pourquoi devrais-je croire le tiers ?

- Confiance
- Transparence
- Méthodes communes

Transparence certifiée



Transparence certifiée

- Fournir des informations sur les plateformes informatiques elles-mêmes, y compris des détails spécifiques sur la façon dont la transparence computationnelle est prise en charge.
- Empaqueter et signer les artefacts résultants ainsi que les enregistrements de leur exécution en utilisant un format standard.

Applications

- Limor, R-squared, cascad, Banque mondiale !
- FSRDC ? IRS ?
- Métadonnées ?

Footnotes

1.

Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014 https://www.force11.org/group/joint-declaration-data-citation-principles-final

2.

National Academies of Sciences, Engineering, and Medicine. 2019. ***Reproducibility and Replicability in Science***. Washington, DC: The National Academies Press. https://doi.org/10.17226/25303.