

1. [Accueil](#)
2. Entity Print

Document de travail

[Économie](#)

# La reproductibilité de la recherche en économie : un cas d'étude

10 Décembre 2021

Auteurs : [Sylvérie Herbert](#), Hautahi Kingi, Flavio Stanchi, Lars Vilhuber

**Document de travail n°853.** Étant donné l'importance de la reproductibilité pour l'éthique scientifique, de plus en plus de revues ont poussé à la transparence de la recherche par le biais de politiques de disponibilité des données. Si l'introduction et la mise en œuvre de telles politiques de données améliorent la disponibilité du code et des données des chercheurs, quel est l'impact sur la reproductibilité? Nous décrivons et présentons les résultats d'un vaste exercice de reproduction dans lequel nous évaluons la reproductibilité des articles de recherche publiés dans l'*American Economic Journal: Applied Economics*, qui a mis en place une politique de disponibilité des données depuis 2005. Notre taux de succès de réplication est relativement modéré, avec 37,78 % de tentatives de réplication réussies. 68 des 162 tentatives de réplication admissibles ont réussi à reproduire l'analyse de l'article (41,98 %), sous réserve de données non confidentielles. 69 autres tentatives (42,59 %) ont abouti à une réplication partielle. Un total de 98 sur 303 (32,34 %) reposaient sur des données confidentielles ou propriétaires, et n'étaient donc pas reproductibles par ce projet. Nous effectuons également plusieurs analyses bibliométriques des articles reproductibles par rapport aux articles non reproductibles et montrons que les articles reproductibles n'apportent pas de bonus de citation aux auteurs.

## [Proportion d'articles avec des données confidentielles et manquantes et ratio de reproductibilité](#)

Comme pour de nombreuses disciplines scientifiques, la transparence et l'accessibilité sont essentielles à la crédibilité de la recherche économique. Cela est d'autant plus vrai que la recherche économique éclaire les décisions de politique économique. En effet, ces dernières années, gouvernements et institutions politiques ont poussé à ce que les politiques publiques soient basées sur des études scientifiques. Ils s'efforcent de fonder leurs politiques sur la recherche académique, d'où la nécessité que les publications scientifiques en économie soient dignes de confiance pour que les politiques soient crédibles. Et pour être digne de confiance, la recherche doit être reproductible.

Le processus d'évaluation par les pairs (*peer-review*) dans les revues économiques garantit la publication de recherches originales et fournit un label de "haute qualité". Toutefois, rédacteurs et évaluateurs n'ont pas l'obligation de vérifier que "les mêmes données et le même code donnent les mêmes résultats". Par conséquent, le processus d'évaluation des articles ne fournit aucun retour sur la reproductibilité. Afin d'améliorer la reproductibilité, les revues ont mis en place des politiques de disponibilité des données (DAP), comme *Econometrica* dès 1933. D'autres revues ont suivi, comme le *Journal of Money, Credit and Banking* à la fin des années 1990. La première revue du top 5, l'*American Economic Review*, a introduit un DAP en 2005, suivie par le *Quarterly Journal of Economics* en 2016. En 2017, seulement 54 % des 343 revues économiques de l'indice de citation Thomson Reuters Social Science avaient une DAP (Höffler, 2017).

Dans cette étude, nous cherchons à vérifier si les politiques de disponibilité des données avec une application souple, comme celle appliquée par l'AEA en 2005, amènent à la reproductibilité. Bien qu'il n'y ait pas de vérification systématique des codes avec de telles politiques, rendre les codes et les données accessibles au public devrait en théorie améliorer la transparence. En fournissant aux étudiants de premier cycle uniquement le code, les données et les informations fournies par les auteurs, nous vérifions s'ils parviennent à reproduire les résultats de l'article.

Tout d'abord, nous obtenons un succès de réplique modéré, malgré une politique de disponibilité des données. Si l'on considère l'échantillon d'articles évalués, seuls 25 % ont été répliqués avec succès, ce chiffre passant à 38 % si l'on considère uniquement les articles éligibles (ceux pour lesquels les auteurs n'ont pas précisé utiliser des données confidentielles). En tenant compte des données non confidentielles, nous trouvons un taux de reproductibilité plus élevé de 43 %. Cette faible reproductibilité est due aux données confidentielles ou manquantes. Même en se limitant aux données non confidentielles, notre analyse montre qu'un nombre substantiel d'articles ont des résultats différents de ceux obtenus en utilisant leurs codes et données. Deuxièmement, nous montrons que même pour les articles reproductibles des changements de code complexes ont été nécessaires pour atteindre des résultats similaires à ceux publiés. Notre analyse suggère que les politiques de disponibilité des données sont nécessaires mais pas suffisantes pour garantir une reproductibilité totale.

Plus important encore, nous montrons que les articles reproduits ne génèrent pas de bonus en termes de citations. L'augmentation des citations ne peut donc peut-être pas être un argument incitatif suffisant pour atteindre des pratiques de documentation et de codage transparentes. Notre analyse appelle donc à une vérification systématique des articles au cours du processus d'arbitrage, afin d'atteindre un équilibre de "bonne reproductibilité".

## **Télécharger l'intégralité de la publication**

[Document de travail n°853. La reproductibilité de la recherche en écon... \(PDF - 1 Mo\)](#)