المنهجيات شبه التجريبية: تقنيات لقياس الأثر دون استخدام العشوائية

عندما لا يكون من الممكن استخدام التوزيع العشوائي، تُعدّ المنهجيات شبه التجريبية بدائل قوية وموثوقة لتقييم الأثر.
يستعرض هذا الموجز ثلاثًا من أكثر هذه المنهجيات استخدامًا: تصميم انقطاع العلاقة بين المتغيرات (RDD)، والسلاسل الزمنية المتقطعة (ITS)، مطابقة بين المشاركين المتشابهين (PSM).
للاطلاع على مزيد من الخلفية والمفاهيم الأساسية، يمكنك الرجوع إلى الموجز التوجيهي الأول بعنوان: أساسيات تقييم الأثر.

هذا الموجز هو جزء من سلسلة مكونة من خمسة موجزات إرشادية أعدّتها مؤسسة الملكة رانيا، استنادًا إلى مواد طوّرها البروفيسور هوارد وايت (مُمثّلًا عن مركز البحوث والتقييم - REC) لدورة تدريبية بعنوان "تقييم الأثر: من التصميم إلى التوسّع". وقد تم تمويل هذه الدورة من قبل مؤسسة الوقف التعليمي (EEF)، ومؤسسة BHP، ومؤسسة الملكة رانيا.

منهجية انقطاع العلاقة بين المتغيرات (RDD)

تعريف:

يقوم تصميم انقطاع المتغير بتقدير التأثيرات السببية من خلال الاستفادة من نقطة قطع محددة مسبقًا في أهلية البرنامج. يعمل المشاركون الذين يقعون مباشرة فوق وتحت هذه العتبة كمجموعتين طبيعيتين للمقارنة، مما يُقلل من تأثير التحيّز في الاختيار.

مثال: مؤشر الفقر والتحويلات النقدية المشروطة

تم تقييم أثر برنامج للتحويلات النقدية استنادًا إلى درجة فقر محددة مسبقًا.

تم استبعاد الأسر التي تجاوزت الحد من الاستفادة، مما شكّل مجموعة ضابطة طبيعية.

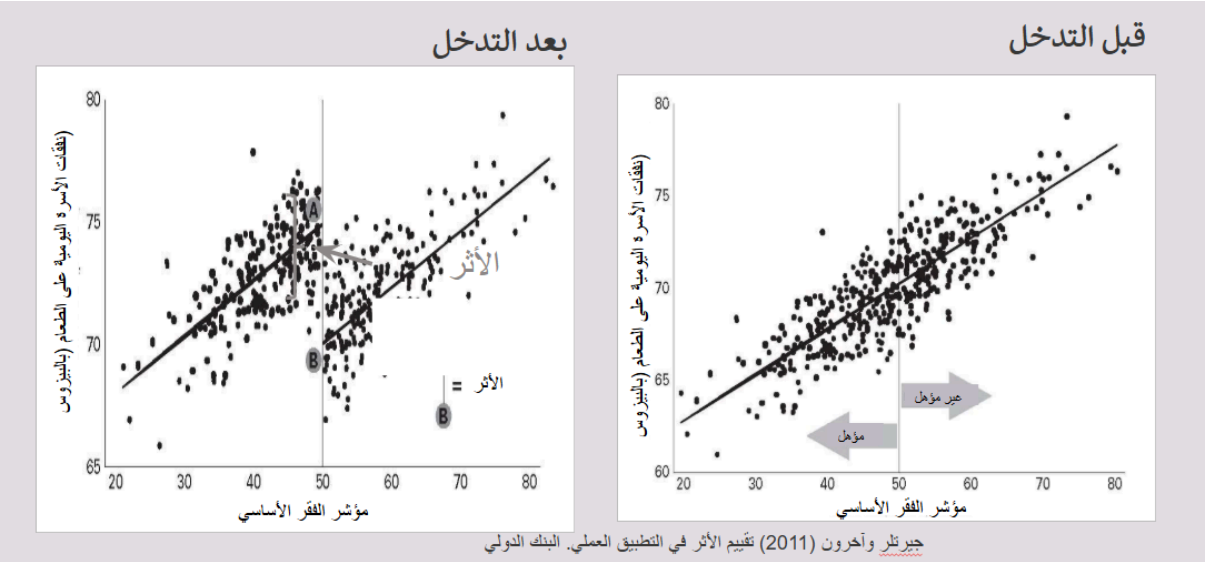
النتائج: لوحظت زيادات ملحوظة في معدلات الالتحاق بالمدارس وتحسّن في تغذية الأطفال، خاصة بين الفتيات.

الرسم البياني أدناه يُبيّن العلاقة بين إنفاق الأسرة على الغذاء ومؤشر الفقر الأساسي. تشير القيم الأعلى للمؤشر إلى كون الأسرة أقل فقرًا، وبالتالي فإن من هم فوق الحد الفاصل لا يكونون مؤهلين للحصول على الدعم.

وكما هو متوقّع، تظهر علاقة وثيقة بين مؤشر الفقر وإنفاق الغذاء. بعد تنفيذ التدخل، تتحسن أوضاع الأسر الواقعة تحت الحد الفاصل، ويزداد إنفاقها على الغذاء.

يُقاس أثر التدخل من خلال حجم "الانقطاع" أو الكسر الهيكلي في العلاقة، ويتم حسابه باستخدام خط انحدار قبل وبعد العتبة.

الشكل 1: الخصائص الأساسية: RDD قبل وبعد التدخل.



جيرتلر وآخرون (2011) تقييم الأثر في التطبيق العملي. البنك الدولي

منهجية السلاسل الزمنية المتقطعة (ITS)

التعريف:

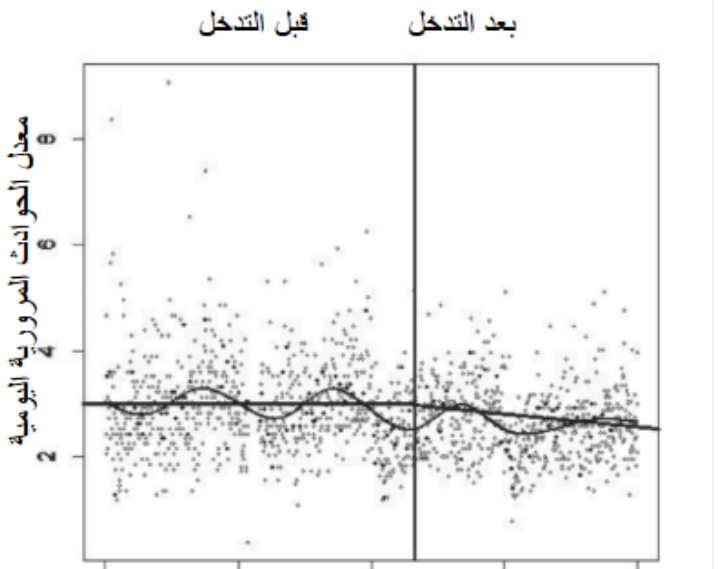
تُستخدم منهجية ITS لتحليل الاتجاهات قبل وبعد تنفيذ التدخل، بهدف تقييم تأثيره على مدى الزمن.

مثال: تجريم القيادة تحت تأثير الكحول في الصين

- تم تقييم أثر قانون جديد على معدلات حوادث الطرق.

النتائج: لم يؤدِّ التدخل إلى انخفاض فوري في الحوادث، لكنه ساهم في تراجع تدريجي مع مرور الوقت، مما يشير إلى تغيّرات سلوكية ناتجة عن تطبيق القانون وزيادة الوعي المجتمعي.

الشكل 2: تجريم القيادة تحت تأثير الكحول في الصين: يوضح هذا المثال تأثير التغيير في القانون على حوادث المرور على الطرق (RTIs)



المصدر: *Zhao, A., Chen, R., Qi, Y., Chen, A., Chen, X., Liang, Z., ... & Kan, H. (2016)*. تقييم أثر تجريم القيادة تحت تأثير الكحول على الإصابات الناجمة عن حركة المرور على الطرق في قوانغتشو، الصين: دراسة متسلسلة زمنية. مجلة علم الأوبئة, *26*(8)، *433-439*.

منهجية المطابقة على أساس درجة الاحتمال (Propensity Score Matching - PSM)

عندما لا يكون التوزيع العشوائي ممكنًا، تُوفّر **المطابقة على أساس درجة الاحتمال** (PSM) بديلًا قويًا لبناء **مجموعة مقارنة** تتشابه في الخصائص القابلة للملاحظة مع المجموعة المستفيدة من التدخل.

الهدف بسيط: التأكد من أن الأفراد الذين تتم مقارنتهم **متشابهون إحصائيًا** (أي يتشاركون في نفس المتوسطات للخصائص القابلة للقياس)، مع اختلاف وحيد هو **تلقي التدخل من عدمه**.

مثال: تقييم برنامج منح دراسية يُقدّم دعمًا ماليًا لطلاب من أسر منخفضة الدخل.

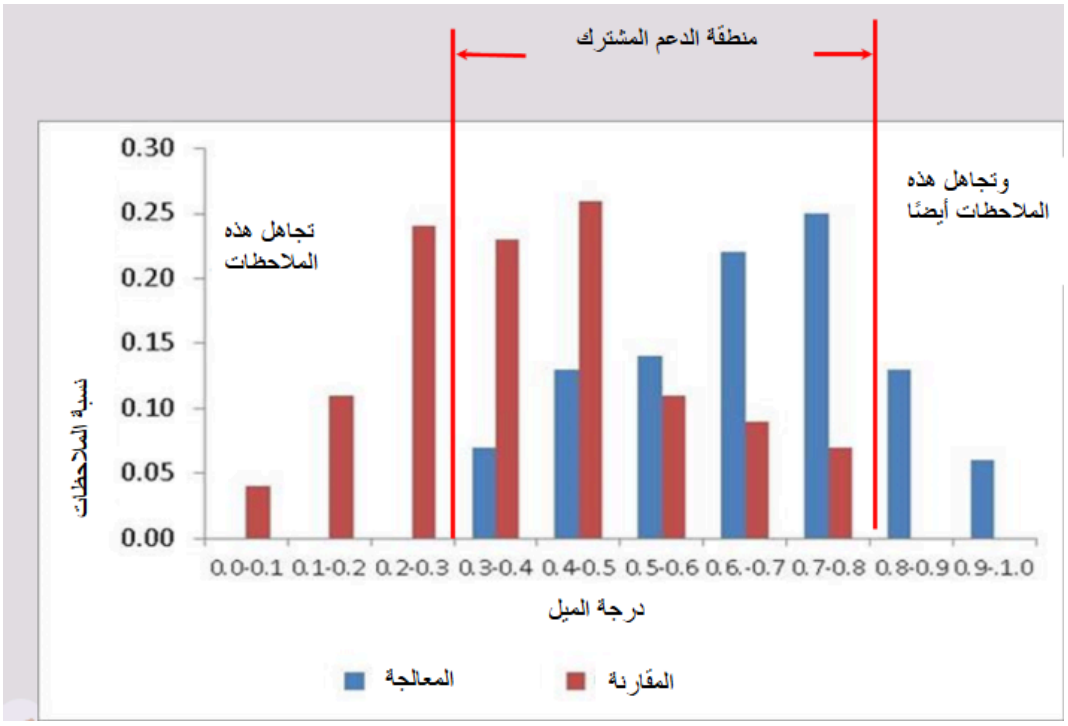
لو قمنا بمقارنة أداء الطلاب الحاصلين على المنحة مع من لم يحصلوا عليها بشكل مباشر، قد تكون النتائج مضلّلة، لأن الطلاب الذين يتقدّمون للمنح عادة ما يكونون أكثر دافعية أو تفوقًا من غيرهم.

تُساعد PSM في الحد من هذا التحيّز، من خلال **حساب احتمال الحصول على المنحة** لكل طالب بناءً على خصائص قابلة للملاحظة، مثل دخل الأسرة، الأداء الأكاديمي السابق، ومستوى تعليم الوالدين.
ومن خلال **مطابقة الطلاب الذين لديهم درجات احتمالية متقاربة**، يتم تكوين مجموعة مقارنة متوازنة، مما يسمح **بعزل أثر المنحة بحد ذاته**.

الخطوات الرئيسية في تنفيذ مطابقة بين المشاركين المتشابهين:

1. تحديد التدخل وتحديد معايير الأهلية.
2. تحديد المتغيرات الأساسية التي تؤثر على احتمالية المشاركة في التدخل.
3. تقدير درجات الميل باستخدام الانحدار اللوجستي أو الاحتمالي.
4. إنشاء منطقة الدعم المشترك وتجاهل الملاحظات التي لا مثيل لها.
5. قم بإجراء نهج المقارنة لأقرب جار.
6. مقارنة النتائج بين المجموعات المتطابقة.
7. إجراء تحليل الحساسية (المتانة) لاختبار قوة النتائج.

الشكل 3: تمثيل لمنطقة الدعم المشترك



رغم أن PSM أداة قوية، إلا أن لها **قيودًا**:

- تفترض أن **جميع المتغيرات المؤثرة** يمكن ملاحظتها وقياسها. وإذا كانت هناك خصائص غير مرئية (مثل الدافعية الذاتية)، فقد يستمر التحيّز.

- لذلك، يُستحسن غالبًا **دمج PSM مع أساليب أخرى** مثل تحليل الفرق في الفرق (Difference-in-Differences – DiD) أو تحليل المتغيرات الأدواتية (Instrumental Variable – IV)، لتعزيز موثوقية النتائج.

الجدول 1: مقارنة بين منهجيات التقييم شبه التجريبية – نقاط القوة والقيود

تصميم	نقاط القوة	القيود
تصميم انقطاع العلاقة بين المتغيرات	صلاحية داخلية قوية، تحاكي العشوائية	صلاحية خارجية محدودة خارج منطقة العتبة. وقد يصعب الحصول على حجم عينة كافٍ، خاصة أنه يقتصر على الحالات القريبة من الحد الفاصل.
السلاسل الزمنية المتقطعة	يأخذ في الاعتبار الاتجاهات على المدى الطويل، ولا يتطلب وجود مجموعة ضابطة	عرضة للعوامل المتداخلة، خصوصًا إذا تزامن التدخل مع سياسات أخرى قد تؤثر على النتائج.
مطابقة بين المشاركين المتشابهين	ينشئ مجموعة مقارنة متوازنة باستخدام خصائص يمكن ملاحظتها، ويخفف من تحيز الاختيار	لا يعالج سوى الخصائص القابلة للملاحظة؛ وقد تستمر الانحرافات الناتجة عن فروقات غير مرئية مثل الدافعية. وقد يكون أكثر ملاءمة في السياقات التي تتوفّر فيها بيانات غنية.

لمحة سريعة: متى نستخدم هذه المنهجيات؟

- **تصميم انقطاع المتغير (RDD): عند وجود حد أهلية صارم يُحدّد من يحق له الاستفادة من التدخل.**
- **السلاسل الزمنية المتقطعة (ITS): عند توفر بيانات زمنية قبل وبعد تنفيذ سياسة أو برنامج.**
- **المطابقة على أساس درجة الاحتمال (PSM): عند تعذّر استخدام التوزيع العشوائي، لكن تتوفر بيانات قوية قبل التدخل.**

لمن يرغب في تطبيق هذه المنهجيات أو تعميق فهمه لها، توفّر الموارد التالية مراجع تجمع بين الوضوح الفني وسهولة الوصول.

المزيد من القراءة

Baird, S., Ferreira, F.H.G., Özler, B. and Woolcock, M. (2013), Relative Effectiveness of Conditional and Unconditional Cash Transfers for Schooling Outcomes in Developing Countries: A Systematic Review. Campbell Systematic Reviews, 9: 1-124. https://doi.org/10.4073/csr.2013.8

Carvalho, S., & White, H. (2004). Theory-based evaluation: the case of social funds. *American journal of evaluation*, *25*(2), 141-160.

Davies R, 2018, Representing Theories of Change: A Technical Challenge with Evaluation Consequences, CEDIL Inception Paper 15: London

Vajja, A., & White, H. (2008). Can the World Bank build social capital? The experience of social funds in Malawi and Zambia. *The Journal of Development Studies*, *44*(8), 1145-1168.

Dickinson-Craig, E., Turbat, T., Hemming, K., Pope, F. D., Bartington, S. E., Anjaa, S., ... & Manaseki-Holland, S. (2025). Improving Air Quality and Childhood Respiratory Health in Mongolia: The Impact of the Raw Coal Ban and COVID-19 Restrictions—An Interrupted Time-Series Analysis. *Atmosphere*, *16*(1), 46.

Zhao, A., Chen, R., Qi, Y., Chen, A., Chen, X., Liang, Z., ... & Kan, H. (2016). Evaluating the impact of criminalizing drunk driving on road-traffic injuries in Guangzhou, China: a time-series study. *Journal of epidemiology*, *26*(8), 433-439.

White, H., & S. Sabarwal (2014). Quasi-experimental Design and Methods, Methodological Briefs: Impact Evaluation 8, UNICEF Office of Research, Florence. https://www.betterevaluation.org/sites/default/files/Quasi-Experimental_Design_and_Methods_ENG.pdf

White, H., & Raitzer, D. A. (2017). *Impact evaluation of development interventions: A practical guide*. Asian Development Bank.

Ikira, M., & Ezzrari, A. (2021). Evaluating the impact of conditional cash transfer programs: Evidence from Morocco. *American Journal of Educational Research*, *9*(5), 320-329. https://users.nber.org/~rdehejia/!@$DRM/Lecture%2001/papers%20to%20discuss%20-%20Research%20vs%20Evlaution/CCTs%20education-9-5-10.pdf