ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Graham, Daniel J.

Working Paper

Inférence causale pour l'évaluation ex post des interventions dans les transports

International Transport Forum Discussion Paper, No. 2014-13

Provided in Cooperation with:
International Transport Forum (ITF), OECD

Suggested Citation: Graham, Daniel J. (2014) : Inférence causale pour l'évaluation ex post des interventions dans les transports, International Transport Forum Discussion Paper, No. 2014-13, Organisation for Economic Co-operation and Development (OECD), International Transport Forum, Paris

This Version is available at:
https://hdl.handle.net/10419/109188

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Forum International des Transports

Inférence causale pour l'évaluation ex post des interventions dans les transports

13

Document de référence
2014 • 13

Daniel J. Graham
Imperial College, Londres,
Royaume-Uni

OCDE

Forum International des Transports

Inférence causale pour l'évaluation ex post des interventions dans les transports

Document de Référence N° 2014-13

Préparé pour la Table Ronde :
L'Évaluation ex post des investissements et interventions publiques dans les transports

(15-16 septembre 2014, OCDE, Paris)

Daniel J. GRAHAM

Department of Civil Engineering
Imperial College, Londres
Royaume-Uni

Novembre 2014

OCDE

FORUM INTERNATIONAL DES TRANSPORTS

Le Forum International des Transports, lié à l'OCDE, est une organisation intergouvernementale comprenant 54 pays membres. Le Forum mène une analyse politique stratégique dans le domaine des transports avec l'ambition d'aider à façonner l'agenda politique mondial des transports, et de veiller à ce qu'il contribue à la croissance économique, la protection de l'environnement, la cohésion sociale et la préservation de la vie humaine et du bien-être. Le Forum International des Transports organise un sommet ministériel annuel avec des décideurs du monde des affaires, des représentants clés de la société civile ainsi que des chercheurs éminents.

Le Forum International des Transports a été créé par une Déclaration du Conseil des Ministres de la CEMT (Conférence Européenne des Ministres des Transports) lors de la session ministérielle de mai 2006. Il est établi sur la base juridique du Protocole de la CEMT signé à Bruxelles le 17 octobre 1953 ainsi que des instruments juridiques appropriés de l'OCDE.

Les pays membres du Forum sont les suivants : Albanie, Allemagne, Arménie, Australie, Autriche, Azerbaïdjan, Bélarus, Belgique, Bosnie-Herzégovine, Bulgarie, Canada, Chili, République populaire de Chine, Corée, Croatie, Danemark, Espagne, Estonie, États-Unis, Ex-République yougoslave de Macédoine, Finlande, France, Géorgie, Grèce, Hongrie, Inde, Irlande, Islande, Italie, Japon, Lettonie, Liechtenstein, Lituanie, Luxembourg, Malte, Mexique, République de Moldova, Monténégro, Norvège, Nouvelle-Zélande, Pays-Bas, Pologne, Portugal, République tchèque, Roumanie, Royaume-Uni, Fédération de Russie, Serbie, République slovaque, Slovénie, Suède, Suisse, Turquie, Ukraine.

Le Centre de Recherche du Forum International des Transports recueille des statistiques et mène des programmes coopératifs de recherche couvrant tous les modes de transport. Ses résultats sont largement disséminés et aident la formulation des politiques dans les pays membres et apportent également des contributions au sommet annuel.

Documents de référence

La série des documents de référence du Forum International des Transports rend les recherches menées par le Centre de Recherche sur les transports ou entreprises à sa demande accessibles aux chercheurs et professionnels du transport. L'objectif est de contribuer tant à la compréhension du secteur des transports qu'à l'élaboration des politiques de transport.

Les documents de référence du FIT ne doivent pas être présentés comme exprimant les vues officielles du FIT ou de ses pays membres. Les opinions exprimées et les arguments employés sont ceux des auteurs.

Les documents de référence exposent des résultats préliminaires ou des travaux de recherche en cours menés par l'auteur/les auteurs et sont publiés pour stimuler le débat sur un large éventail de questions sur lesquelles le FIT travaille. Les commentaires sur les documents de référence sont bienvenus et peuvent être adressés à : Forum International des Transports/OCDE, 2 rue André Pascal, 75775 Paris Cedex 16, France.

Pour de plus amples renseignements sur les documents de référence et les autres activités du CCRT, veuillez envoyer un courriel à : itf.contact@oecd.org

Les documents de référence peuvent être déchargés à l'adresse suivante : www.internationaltransportforum.org/jtrc/DiscussionPapers/jtrcpapers.html

Le site Web du Forum International des Transports est : www.internationaltransportforum.org

Ce document et toute carte qu'il peut comprendre ne préjugent en rien du statut de tout territoire, de la souveraineté s'exerçant sur ce dernier, du tracé des frontières et limites internationales, et du nom de tout territoire, ville ou région.

Table des Matières

RÉSUMÉ....4

1. INTRODUCTION....4

2. L'ÉVALUATION *EX POST* SOUS LA FORME D'UN PROBLÈME D'ESTIMATION DES EFFETS DE TRAITEMENT....6

Difficultés que pose l'estimation de l'effet causal des traitements sur les résultats....6
Données manquantes....8
Assignation non aléatoire et problème de la confusion....9
Identification des effets causals au moyen du modèle des résultats potentiels....10

3. MÉTHODES CAUSALES POUR L'ESTIMATION DES EFFETS DE TRAITEMENT....13

Estimation des effets de traitement avec ignorabilité....13

4. APPLICATIONS....16

Évaluation *ex post* des augmentations de la capacité routière urbaine aux États-Unis....16
Évaluation *ex post* des impacts économiques régionaux du train à grande vitesse en Espagne....21

5. CONCLUSIONS....23

BIBLIOGRAPHIE....24

Résumé

On examine dans ce document les méthodes visant à tirer des inférences causales de données non expérimentales et on montre comment elles peuvent s'appliquer à l'évaluation ex post des interventions dans les transports. En particulier, on considère les principes sous-jacents des techniques employées pour l'estimation des effets de traitement avec des traitements assignés de manière non aléatoire. Ces techniques ont pour but de quantifier les changements qui se sont produits du fait d'une intervention (« traitement ») explicite. L'idée est ici que les interventions dans les transports se caractérisent typiquement par une assignation non aléatoire et que les questions clés pour une réaliser une bonne évaluation ex post résident dans l'identification et la correction des facteurs de confusion. Contrairement aux approches classiques de l'appréciation ex ante, un avantage majeur des méthodes causales statistiques est qu'on peut les appliquer sans poser de fortes hypothèses théoriques a priori. On présente dans ce document des exemples empiriques de l'utilisation des techniques causales pour l'évaluation des augmentations de capacité du réseau routier dans les grandes villes des États-Unis et pour les investissements dans le train à grande vitesse en Espagne.

Mots-clés : Inférence causale ; effet de traitement ; score de propension ; évaluation ex post.

1. Introduction

L'évaluation ex ante des transports a de solides racines théoriques et empiriques dans les calculs d'analyse coûts-avantages (ACA) sur la base de la rente du consommateur. Des travaux récents sur les « avantages économiques élargis » ont étendu l'ACA « standard » de manière à y incorporer certaines externalités clés et des formes de concurrence imparfaite, là encore sur la base d'éléments théoriques et empiriques clairement établis. L'ACA offre une approche familière et bien comprise couramment utilisée par les agents publics, les spécialistes des transports et les universitaires.

Une attention bien moindre a été consacrée à l'évaluation ex post des transports aussi bien en théorie qu'en pratique. C'est peut-être parce qu'en général nous nous intéressons plus à prédire le degré de succès de nos investissements futurs qu'à évaluer avec quelle habileté nous avons alloué les ressources dans le passé. Pourtant, si nous voulons bien comprendre les impacts qu'auront les interventions dans les transports, une bonne base de référence peut être établie en étudiant les interventions précédentes et leur degré de succès au regard de certaines métriques définies comme il convient.

Une façon de le faire consiste simplement à ré exécuter les calculs d'ACA *ex ante* quelque temps après l'achèvement du projet en utilisant les valeurs observées, au lieu des valeurs prédites. Cela peut fournir des informations utiles à la fois sur les impacts du projet lui-même et sur l'exactitude avec laquelle l'ACA *ex ante* a été capable de prédire les coûts et avantages du dispositif. Ces calculs

restent toutefois générés à l'intérieur du cadre théorique de l'ACA qui pose un certain nombre d'hypothèses simplificatrices assez restrictives. Une autre approche consiste à appliquer des modèles statistiques aux données observées avant et après les interventions dans les transports et essayer d'estimer les impacts qui ont été causés par l'intervention. Ces approches statistiques, que nous désignons par le terme de méthodes d'inférence causale, s'appuient plus sur des méthodes empiriques que sur la théorie économique, mais elles ont leurs propres hypothèses et propriétés qui doivent être satisfaites pour générer une inférence causale valide.

Dans le présent document, nous examinons les approches statistiques couramment utilisées dans diverses disciplines scientifiques pour inférer des relations cause-effets à partir des données observationnelles. Nous pensons qu'un cadre d'inférence causale est très approprié à l'évaluation *ex post* parce qu'il est spécifiquement conçu pour estimer les effets produits quand des « traitements » sont assignés de manière non aléatoire, comme c'est le cas pour la plupart des interventions dans le domaine des transports. La conséquence essentielle de l'assignation non aléatoire est que l'effet du traitement subit une « confusion », impliquant que les unités qui reçoivent le traitement (ou une dose particulière du traitement) peuvent différer de manière systématique des unités ayant un statut différent au regard du traitement. L'analyse causale a pour objectif d'estimer l'effet moyen du traitement (intervention) net de confusion ou, autrement dit, de révéler l'effet causal marginal. Nous désignons cela par le terme de problème d'estimation des effets de traitement et c'est dans ce contexte que nous examinons les méthodes possibles pour l'évaluation ex post.

Ce document est structuré comme suit. On examine dans la section 2 l'évaluation ex post comme un problème d'estimation des effets de traitement dans le cadre du « modèle des résultats potentiels » pour l'inférence causale. On décrit les implications de l'assignation non aléatoire du traitement et le problème de la confusion et on présente ensuite les hypothèses clés requises pour une inférence causale valide. La section 3 expose les stratégies pour une estimation convergente de l'effet de traitement. On étudie l'identification des effets causals dans des conditions « d'ignorabilité » par l'ajustement de covariables, l'ajustement du score de propension et des méthodes doublement robustes. On examine ensuite deux approches couramment utilisées quand on ne suppose pas l'hypothèse d'ignorabilité. La section 4 présente deux exemples empiriques de l'utilisation de techniques causales pour l'évaluation ex post : le premier évalue les impacts des augmentations de capacité du réseau routier aux États-Unis et l'autre considère les impacts économiques régionaux des investissements dans le train à grande vitesse en Espagne. On tire ensuite des conclusions dans la section finale.

2. L'évaluation *ex post* sous la forme d'un problème d'estimation des effets de traitement

Notre insistance sur la causalité dans le présent document découle de la conviction que la politique des transports a pour préoccupation fondamentale les relations cause-effet. Au Royaume-Uni, par exemple, les questions suivantes ont eu une forte influence, ces dernières années, dans les prises de décision :

- Quel effet la fiscalité des carburants aura-t-elle sur les émissions dans le domaine des transports ?
- Dans quelle proportion les volumes de trafic diminueront-ils dans le cadre d'un péage de congestion ?
- Quelle sera la variation de la demande de déplacements avec l'élévation du niveau de vie ?
- L'investissement dans l'infrastructure de transports stimulera-t-il la productivité de l'économie ?
- Quel sera l'effet des options d'investissement sur les performances du réseau ?

Chacune de ces questions comporte une relation cause-effet et le but sous-jacent des politiques est d'essayer de façonner les résultats futurs par l'intervention publique. Pour les décideurs, il s'agit de savoir quel impact, ou résultat, auront les interventions envisagées.

L'évaluation ex post peut aider à y répondre. En appliquant des modèles statistiques aux données historiques, on peut essayer de saisir les relations clés et d'évaluer l'effet d'interventions passées sur des résultats définis. Toutefois, l'interprétation causale des données observées se heurte à deux problèmes fondamentaux. Premièrement, nous n'observons que ce qui a eu lieu réellement et non ce qui aurait eu lieu si on était intervenu de manière différente. Deuxièmement, les interventions que nous faisons sont rarement assignées de manière aléatoire et cette assignation non aléatoire obscurcit les relations cause-effet. Conjointement, ces deux problèmes signifient en fait que nous ne possédons pas d'éléments expérimentaux pouvant servir de base aux décisions publiques.

Dans cette section, on décrit le modèle des résultats potentiels pour l'inférence causale, que l'on peut utiliser pour obtenir une interprétation causale des données observationnelles en l'absence de conditions expérimentales. On examine les caractéristiques définissant cette approche et on montre comment elle peut permettre d'inférer des relations cause-effet pour les interventions dans les transports.

Difficultés que pose l'estimation de l'effet causal des traitements sur les résultats

Trois composantes clés sont à prendre en compte quand on analyse les relations cause-effet à partir des données observées : l'intervention (ou traitement) à étudier, le résultat considéré et les

caractéristiques pertinentes des unités d'observation. Pour l'analyse ex post, nous nous intéressons fondamentalement à l'effet qu'une intervention (ou un ensemble d'interventions) dans le domaine des transports a sur un résultat. Nous souhaitons éventuellement savoir ce qu'aurait été le résultat si l'intervention n'avait pas eu lieu ou si une intervention différente avait été mise en œuvre.

Les résultats à considérer peuvent concerner les conditions de trafic (vitesses, flux, sécurité, congestion), des caractéristiques économiques (production, productivité, croissance), la part modale, les conséquences environnementales, des préoccupations sociales, etc.

Aux fins de l'analyse, on peut considérer une intervention dans le système de transports comme une réalisation observée de variables aléatoires dont la manipulation produit différents résultats. Nous désignons ces variables aléatoires par le terme de « traitements », définis dans le sens le plus large de manière à couvrir tout « régime » qui peut être manipulé pour produire un certain effet. Par exemple, un traitement peut comporter la construction d'une nouvelle liaison, l'imposition de limites de vitesse, des changements de prix des transports, des changements de fréquence ou de qualité de service, l'attribution de subventions, etc. Les variables de traitement peuvent être binaires, plurivalentes ou continues. Le Tableau 1 donne des exemples d'interventions dans les transports catégorisées en tant que variables de traitement.

Tableau 1. **Interventions dans les transports catégorisées en tant que variables de traitement**

Binaire	**Plurivalente**	**Continue**
Route à péage/sans péage	Fréquence de service	Capacité de réseau
Présence de radar de vitesse	Limite de vitesse	Longueur de trajets en site propre
Désignation de zone à 30 km/h	Nombre de wagons par train	Densité de carrefours
Heures de pointe / heures creuses	Nombre d'itinéraires par O-D	Accessibilité
Piétonisation / non-piétonisation	Nombre de nœuds de réseau	Taux de taxes / subventionnement

Nous nous intéressons à l'effet du traitement sur le résultat, mais nous savons aussi que les unités étudiées n'auront probablement pas des caractéristiques homogènes, ce qui peut avoir une influence dans la relation fondamentale considérée. Suivant le but de l'analyse, et les données disponibles, les unités peuvent être des dispositifs de transport particuliers, des liaisons de réseau, des personnes, des ménages, des entreprises, des zones géographiques, des villes, etc.

Nous voulons estimer les effets causals au moyen de données représentant ces trois composantes. Nous définissons $z_i = (y_i, d_i, x_i)$, $i = 1, ..., n$, comme un vecteur aléatoire de données observées où, pour la i-ème unité d'observation, y_i désigne un résultat (ou réponse), d_i le traitement (ou exposition) reçu, et x_i un vecteur de co-variables avant traitement. Comme mentionné précédemment, le traitement peut être binaire (c'est-à-dire $D \in \{0, 1\}$), plurivalent, où la dose d peut prendre des valeurs dans m catégories, $D \equiv (d_0, d_1, ..., d_m)$, ou continu avec une dose d qui prend des valeurs dans $D \subseteq R$.

Nous voulons estimer l'effet des traitements sur les résultats. À cette fin, nous utilisons le modèle des résultats potentiels pour l'inférence causale, qui a été présenté pour la première fois par Rubin pour les traitements binaires dans une série d'articles au cours des années 70 (par exemple, Rubin 1973a,b, 1974, 1977, 1978), bien que cet auteur voie dans les travaux antérieurs de Fisher (1935) et de Neyman (1923) des précurseurs de son approche. Le modèle des résultats potentiels définit les conditions dans lesquelles on peut estimer les effets causals à partir des données observées. Deux questions fondamentales interviennent dans l'approche des résultats potentiels.

Données manquantes

Dans l'idéal, nous calculerions l'effet de chaque traitement de manière individuelle (unité par unité). Ainsi, pour l'unité i et le traitement binaire D ∈ {0, 1} nous pouvons définir deux résultats potentiels : $Y_i(0)$ si $D_i = 0$, et $Y_i(1)$ si $D_i = 1$. L'effet causal individuel (ECI) du traitement est alors défini comme :

$$\tau_i = [Y_i(1) - Y_i(0)].$$

Pour un traitement plurivalent ou continu, on peut définir un résultat potentiel $Y_i(d)$ associé à chaque dose de traitement d, $Y_i = \{Y_i(d) : d \in D\}$ désignant l'ensemble complet des résultats potentiels. Les ECI correspondants seraient alors

$$\tau_i = [Y_i(d) - Y_i(0)],$$

c'est-à-dire la différence entre le résultat avec assignation à la dose d et celui en absence de traitement.

Un problème fondamental pour l'inférence causale tient toutefois dans le fait que les données disponibles pour l'estimation ne révèlent que les résultats effectifs et non les résultats potentiels. Pour un traitement binaire, nous observons

$$Y_i = Y_i(1)\, I_1(D_i) + Y_i(0)[1 - I_1(D_i)],$$

où $I_1(D_i)$ est la fonction indicatrice pour la réception du traitement, mais nous n'observons pas la densité conjointe, f $[Y_i(0), Y_i(1)]$, étant donné que les deux résultats ne se produisent jamais simultanément. Pour les traitements plurivalents ou continus, nous n'observons que $Y_i(D_i)$, et les résultats à tous les autres niveaux, d /= D_i, sont inobservés et nous les appelons résultats contrefactuels.

Ainsi, la difficulté est que les données observées ne fournissent pas assez d'information pour évaluer les ECI parce que nous n'observons pas les résultats potentiels découlant des assignations de traitement qui sont contraires au fait. Pour Holland (1986), cela constitue un problème d'identification fondamental de l'inférence causale. Une idée clé de l'approche des résultats potentiels est que, si l'on s'attache à estimer les effets causals moyens, au lieu des ECI, il n'est pas alors nécessaire d'observer tous les résultats potentiels.

Les estimandes causals moyens que l'on considère sont les résultats potentiels moyens (RPM) et les effets de traitement moyens (ETM). Pour les traitements binaires, les RPM sont

$$\mu(1) = E[Yi(1)] \text{ et } \mu(0) = E[Yi(0)],$$

et l'ETM est défini par :
$\tau(1) = \mu(1) - \mu(0)$.

Pour les traitements continus et plurivalents, le RPM pour le niveau de traitement d est

$\mu(d) = E[Yi(d)]$,

et l'ETM est

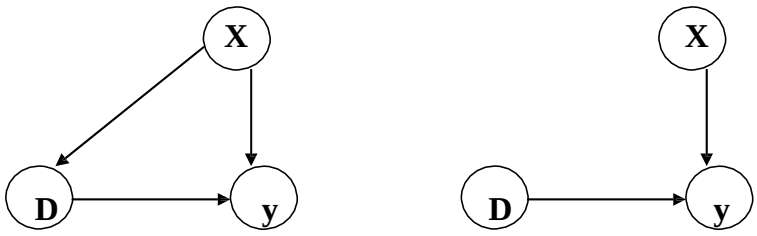
$\tau(d) = \mu(d) - \mu(0)$.

On peut considérer d'autres estimandes causals comme les ETM sur les unités traitées, les effets par quantile, les ETM pour diverses sous-populations, les ETM conditionnés par des covariables et les rapports de côtes et risques relatifs. Dans le présent document, on s'intéresse essentiellement aux RPM et ETM définis ci-dessus.

Assignation non aléatoire et problème de la confusion

Les conditions dans lesquelles on peut utiliser les données observées pour estimer les RPM et les ETM dépendent de manière cruciale du point du savoir si le traitement est assigné ou non de manière aléatoire. Le graphique 1 ci-dessous compare par un schéma les assignations de traitement aléatoires et non aléatoires.

Figure 1. **Graphe orienté acyclique des données observationnelles avec et sans randomisation de l'assignation du traitement**



non aléatoire aléatoire

Dans le cas d'une assignation aléatoire, les caractéristiques X des unités n'ont pas d'influence sur le traitement reçu (c'est-à-dire sur D). En conséquence, les résultats potentiels sont inconditionnellement indépendants du mécanisme d'assignation de traitement. Pour les traitements binaires, la randomisation implique :

$Y_i(0), Y_i(1)) \perp I_1(D_i)$

et pour les traitements plurivalents ou continus

$Y_i(d) \perp I_d(D_i)$ pour tout $d \in D$,

où $I_d(D_i)$ est la fonction indicatrice associée à la réception de la dose d du traitement. Dans le cas d'une assignation aléatoire, il n'y a pas de différences systématiques des caractéristiques entre les

unités traitées ou témoins ou, dans le cas des traitements plurivalents ou continus, entre les unités qui reçoivent différentes doses du traitement. En conséquence, nous pouvons traiter les résultats potentiels inobservés en grande partie comme des données aléatoirement manquantes et on peut construire des estimateurs convergents des ETM pour les traitements binaires et pour les traitements plurivalents ou continus sous la forme :

$$\hat{\tau}(1) = \frac{\sum Y_i \cdot I_1(D_i)}{\sum I_1(D_i)} - \frac{\sum Y_i \cdot (1 - I_1(D_i))}{\sum (1 - I_1(D_i))}$$

et

$$\hat{\tau}(d) = \frac{\sum Y_i \cdot I_d(D_i)}{\sum I_d(D_i)} - \frac{\sum Y_i \cdot I_0(D_i)}{\sum I_0(D_i)}$$

En cas de non-randomisation, au contraire, l'assignation du traitement dépend d'un ensemble de covariables X qui sont elles-mêmes importantes dans la détermination du résultat Y . Ainsi, une partie de l'association entre le traitement et le résultat pourrait être attribuée à X plutôt qu'à D. Dans ces circonstances, nous appelons X des confondeurs et notons que les simples comparaisons de réponses moyennes entre différents groupes de traitement (comme dans les équations 1 et 2) ne révélerons pas en général un effet « causal » parce que les résultats moyens pour les unités traitées ou témoins différeront sans considération de la situation au regard du traitement.

Identification des effets causals au moyen du modèle des résultats potentiels

Bien que l'assignation non aléatoire ait des conséquences pour l'estimation causale, on peut néanmoins obtenir des estimations convergentes des RPM et des ETM au moyen du modèle des résultats potentiels, mais il faut faire un ajustement pour la confusion. Dans cette sous-section du document, nous définissons les conditions dans lesquelles on peut identifier des estimations causales en présence de confusion.

Trois hypothèses clés sont requises pour une estimation valide des RPM et ETM en présence de confusion, à l'intérieur du modèle des résultats potentiels :

1. **Indépendance conditionnelle** – Les résultats potentiels pour l'unité i doivent être conditionnellement indépendants de l'assignation de traitement sachant un ensemble (suffisant) de covariables X_i observées. Pour les traitements binaires, l'hypothèse exige que :

   $$Y_i(0), Y_i(1)) \perp I_1(D_i)|X_i, \quad (3)$$

   et, pour les traitements plurivalents ou continus, Imbens (2000) et Hirano et Imbens (2004) introduisent le concept d'indépendance conditionnelle faible, qui peut s'écrire :

   $$Y_i(d) \perp I_d(D_i)|X_i \text{ pour tout } d \in D. \quad (4)$$

   La différence essentielle entre les hypothèses binaire et non binaire est que, dans ce dernier cas, il doit y avoir, non pas indépendance conjointe de tous les résultats potentiels, mais indépendance conditionnelle pour chaque valeur du traitement (c'est-à-dire par paire).

   Les hypothèses d'indépendance conditionnelle exigent essentiellement que, conditionnellement à un ensemble de covariables avant traitement, l'assignation au traitement ne dépende pas du résultat. Si X_i est suffisant pour que cela soit vérifié, nous pouvons alors en fait imiter,

pour les données observationnelles, l'assignation qui aurait lieu dans un essai comparatif randomisé où le traitement est assigné indépendamment des caractéristiques avant traitement.

2. **Support commun** – le support de la distribution conditionnelle de X_i pour une situation donnée au regard du traitement doit chevaucher celui de X_i pour toute autre situation donnée au regard du traitement. Pour les traitements binaires, cela exige que la probabilité de l'assignation au traitement soit strictement comprise entre 0 et 1

$$0 < \Pr(I_1(D_i) = 1|X_i = x) < 1, \forall\, x. \tag{5}$$

Pour les traitements plurivalents ou continus, nous exigeons un support commun par situation au regard du traitement pour les distributions des covariables à l'intérieur d'une certaine région de dose $C \subseteq D$. Une condition suffisante est que pour tout sous-ensemble de C, disons $A \subseteq C$,

$$\Pr(D_i \in A|X_i = x) > 0, \forall\, x \tag{6}$$

L'idée intuitive sous-jacente à l'hypothèse du support commun, ou chevauchement, est que, si certaines sous-populations observées dans X_i ont une probabilité nulle de recevoir (ou de ne pas recevoir) un traitement, il n'y a pas de sens dans ce cas à parler d'un effet de traitement étant donné que le contrefactuel n'existe pas dans les données observées.

3. **Hypothèse SUTVA** – la relation entre les résultats observés et potentiels doit satisfaire à l'hypothèse SUTVA (Stable Unit Treatment Value Assumption) (voir, par exemple, Rubin, 1978, 1980, 1986, 1990), qui exige que la réponse observée dans une assignation de traitement donnée soit équivalente à la réponse potentielle dans cette assignation. Pour les traitements binaires, nous exigeons donc que

$$Y_i = I_1(D_i)Y_i(1) + (1 - I_1(D_i))Y_i(0) \tag{7}$$

pour tout i = 1, ..., N. Pour les traitements plurivalents ou continus, nous exigeons que

$$Y_i \equiv I_d(D_i)Y_i(d) \tag{8}$$

pour tout $d \in D$, pour tout $Y_i(d) \in Y_i$ et pour i = 1, ..., N.

L'hypothèse SUTVA nécessite que le résultat pour chaque unité soit indépendant de la situation des autres unités au regard du traitement ou, autrement dit, qu'il n'y ait pas d'interférence dans les effets du traitement entre les unités (Cox 1958). Il implique également qu'il n'y ait pas différentes versions du traitement. L'hypothèse d'absence d'interférence est généralement satisfaite quand les unités sont physiquement distinctes et n'ont pas de moyens de contact. Il peut y avoir des violations de l'hypothèse quand la proximité des unités permet un contact et cela constitue une préoccupation particulière pour les applications aux transports.

Les trois hypothèses énoncées ci-dessus, dont Rosenbaum et Rubin (1983) désignent l'ensemble par le terme d'ignorabilité forte, assurent l'identifiabilité des effets causals à partir des données observationnelles. Dans le cas des traitements binaires, l'ETM peut se calculer comme suit,

$$\tau = E_i(Y_i(1) - Y_i(0)) = E_X\,[E_i(Y_i(1)|X_i = x) - E_i(Y_i(0)|X_i = x)] \tag{9a}$$
$$= E_X\,[E_i(Y_i(1)|X_i = x, I_1(D_i) = 1) - E_i(Y_i(0)|X_i = x, I_1(D_i) = 0)] \tag{9b}$$
$$= E_X\,[E_i(Y_i|X_i = x, I_1(D_i) = 1) - E_i(Y_i|X_i = x, I_1(D_i) = 0)]. \tag{9c}$$

L'indépendance conditionnelle justifie l'égalité de (9a) et (9b), l'hypothèse SUTVA permet de remplacer les résultats potentiels par les résultats observés pour obtenir (9c) et le chevauchement garantit que l'ETM de la population dans (9c) est estimable étant donné qu'il existe des unités à la fois dans le groupe traité et dans le groupe non traité. On notera que l'ETM est défini comme une espérance sur les covariables X. Si l'on n'applique pas cette espérance, mais qu'on utilise l'intégrande, on obtient une estimation de l'effet causal de D à l'intérieur des strates de X. Autrement dit, on obtient l'effet de traitement conditionnel, qui est l'effet de traitement moyen pour les unités ayant les caractéristiques X = x. En intégrant par rapport à X, on obtient l'effet causal moyen.

Pour les traitements plurivalents ou continus, le RPM pour une dose donnée --
D = d, $\mu(d) = E[Y_i(d)]$, ou fonction dose-réponse -- se calcule par:

$$E[Y_i(d)] = E_X\,[E(Y_i(d)|X_i)] = E_X\,[E(Y_i(d)|I_d(D_i), X_i)] = E_X\,[E(Y_i|I_d(D_i), X_i)]\,, \qquad (10)$$

où la deuxième égalité découle de l'indépendance conditionnelle, la troisième de la SUTVA et où l'hypothèse de chevauchement garantit que le RPM est estimable étant donné qu'il existe des unités comparables entre les niveaux de traitement.

3. Méthodes causales pour l'estimation des effets de traitement

Les travaux publiés sur les méthodes d'estimation causale forment un vaste ensemble qui continue de croître rapidement. En conséquence, il n'est pas possible de présenter ici un examen véritablement complet de ce domaine et ce serait en fait peu utile étant donné qu'il existe déjà d'excellents inventaires récents, comme ceux de Hernan et Robins (2012) ; Imbens et Wooldridge (2009) ; Tsiatis (2006) ; ou van der Laan et Robins (2003). Nous nous contenterons d'exposer les principes généraux gouvernant la construction des estimateurs. Nous commençons par les méthodes qui supposent l'hypothèse d'ignorabilité et nous considérons ensuite deux approches fréquemment employées quand on pense que cette hypothèse n'est pas satisfaite.

Estimation des effets de traitement avec ignorabilité

En utilisant la notation de Tsiatis et Davidian (2007), nous définissons la densité conjointe des données observées sous la forme : $f_Z(z) = f_{Y|D,X}(y|d, x) f_{D|X}(d|x) f_X(x)$.

En supposant que l'ignorabilité est vérifiée, l'estimation des RPM et des ETM suit généralement une des méthodes suivantes :

a) Ajustement direct des covariables – on laisse $f_{D|X}(d|x)$ et $f_X(x)$ non spécifiées et on postule un modèle pour $E[Y_i|D_i, X_i]$, l'espérance de la densité conditionnelle de la réponse sachant le traitement et les covariables. On utilise typiquement un modèle de régression des résultats (RR) tel que le modèle linéaire généralisé (MLG), le modèle mixte linéaire généralisé (MMLG), le modèle mixte additif généralisé (MMAG) ou d'autres approches sur la base de splines. Les ETM peuvent directement être estimés à partir de ces modèles RR. Cette approche de type régression est communément utilisée dans les analyses de transport.

b) Ajustement du score de propension - on laisse $f_{Y|X}(y|x)$ et $f_X(x)$ non spécifiées mais on suppose un modèle pour $f_{D|X}(d|x)$ et on les utilise pour établir des scores de propension (SP), qui mesurent la probabilité de l'assignation au traitement sachant l'ensemble de covariables avant traitement observées. Un résultat important, dû à Rosenbaum et Rubin (1983), énonce que l'hypothèse d'indépendance conditionnelle (c'est-à-dire les équations 3 et 4) peut être reformulée en remplaçant le vecteur de covariables X_i par le scalaire SP. Rosenbaum et Rubin (1983) ont démontré ce résultat dans le cas des traitements binaires et Imbens (2000) et Hirano et Imbens (2004) généralisent le SP au cas des traitements plurivalents ou continus.

Le SP sert à construire divers estimateurs non paramétriques ou semi-paramétriques par pondération, appariement, stratification, groupage ou régression (pour des détails, voir Imbens et Wooldridge 2009). Un des avantages fondamentaux de l'utilisation du SP est d'éviter la nécessité de conditionner sur un vecteur de covariables potentiellement à un grand nombre de dimensions et c'est cette propriété de réduction des dimensions qui permet la mise en œuvre effective d'estimateurs flexibles. Un autre avantage du SP est qu'il est très efficace pour isoler la région de support commun, tâche qui est difficile quand on utilise de nombreuses covariables (sur cette question, voir Joffe et Rosenbaum 1999).

c) Estimation doublement robuste - on laisse f (x) non spécifiée mais on suppose à la fois un modèle RR et un modèle SP et on construit un estimateur qui combine les deux modèles. On procède habituellement en pondérant ou en augmentant le modèle RR avec des covariables calculées en inversant les SP. La caractéristique essentielle des estimateurs doublement robustes est que les estimations des RPM et ETM sont convergentes et asymptotiquement normales quand le modèle OR ou le modèle SP sont correctement spécifiés, mais sans exiger que les deux soient corrects (par exemple, Robins 2000 ; Robins *et al.*, 2000 ; Robins et Rotnitzky, 2001 ; van der Laan et Robins, 2003 ; Lunceford et Davidian, 2004 ; Bang et Robins, 2005 ; Kang et Schafer, 2007). La raison d'être de l'estimation doublement robuste est que l'analyste a en fait deux chances de spécifier correctement le modèle.

Estimation étant donné une assignation de traitement non ignorable

La validité des approches d'estimation examinées dans la sous-section précédente suppose que l'hypothèse d'ignorabilité est vérifiée. Quand celle-ci n'est plus tenable, il faut considérer d'autres méthodes. Dans la présente sous-section, nous examinons deux estimateurs fréquemment utilisés quand on ne suppose plus l'ignorabilité : les différences de différences et les variables instrumentales.

a) Différences de différences – un problème courant dans l'identification des effets causals est qu'il peut y avoir des différences inobservées entre les unités traitées et non traitées qui influent sur les résultats potentiels et interviennent aussi dans l'assignation du traitement. En outre, on peut soupçonner que les tendances temporelles influent sur la variable de résultat en raison d'événements sans lien avec le traitement.

L'estimateur des différences de différences (DDD) répond à ces sources de biais potentielles en utilisant des informations à la fois pour les groupes traité et témoin et dans les périodes avant et après traitement. L'estimateur des DDD approxime :

$$\tau_{DID} = \{E[Y_i(1)|D=1] - E[Y_i(1)|D=0]\} - \{E[Y_i(0)|D=1] - E[Y_i(0)|D=0]\} \quad (11)$$

La « double différence » de l'estimateur des DDD supprime deux sources de biais potentielles. Premièrement, elle élimine les biais dans les comparaisons de seconde période entre les groupes traité et témoin qui pourraient émaner des caractéristiques invariantes dans le temps. Deuxièmement, elle corrige les biais variant dans le temps dans les comparaisons pour le groupe traité au cours du temps qui pourraient être attribuables aux tendances temporelles sans lien avec le traitement.

Il importe de noter deux limitations potentielles concernant l'approche par les DDD. Premièrement, elle s'appuie sur l'hypothèse forte d'identification suivant laquelle les résultats moyens pour les groupes traités et témoin auraient suivi des trajectoires parallèles au cours du temps en l'absence de traitement.

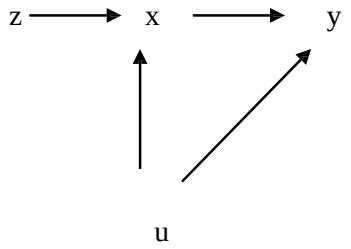
Deuxièmement, le modèle est sensible à la spécification des erreurs et, en particulier, il a été montré que l'existence d'une corrélation à l'intérieur de groupes ou temporelle peut nuire aux performances de l'estimateur des DDD (Bertrand *et al.*, 2004).

b) Variables instrumentales (VI). L'estimateur des VI est bien connu et très utilisé, et pour cette raison nous ne présentons pas ici un examen détaillé. Les principes fondamentaux de l'estimation par les VI sont les suivants :

1. Trouver un ensemble d'instruments qui sont exogènes et fortement corrélés avec les covariables.
2. Utiliser ces instruments pour imposer l'orthogonalité entre le terme d'erreur et la matrice transformée des instruments.

Les relations supposées dans l'estimation VI sont présentées par le schéma ci-dessous dans le contexte du modèle de régression linéaire $y = X\beta + u$ avec la matrice d'instruments Z.

Figure 2. **Relations dans l'estimation par les variables instrumentales**



Les caractéristiques définissant le modèle VI sont les suivantes : les changements de z sont associés aux changements de x, mais n'induisent pas de changements de y autres qu'à travers x ; z est causalement associé à x mais certainement pas à u ; z ne figurerait pas dans le modèle de régression pour y.

Une méthode couramment utilisée pour obtenir des estimations VI est celle des doubles moindres carrés (DMC) :

1. On régresse chaque colonne de X sur la matrice des instruments Z ;
2. On régresse y sur les valeurs prédites découlant de la première phase.

La méthode des VI peut servir à établir des effets causals dans le cas d'une assignation de traitement non ignorable et elle est particulièrement utile quand il y a endogénéité par causalité bidirectionnelle. Toutefois, il est crucial que les deux hypothèses clés d'exogénéité et de pertinence soient satisfaites et, dans la pratique, ces instruments peuvent être difficiles à trouver. Quand les instruments ne sont que faiblement corrélés aux régresseurs endogènes, ou quand les instruments eux-mêmes sont corrélés au terme d'erreur, la méthode des VI peut produire des estimations biaisées et non convergentes. Ce problème est encore compliqué par le fait que la statistique diagnostique existante ne donne pas des moyens parfaits de détecter une spécification inadéquate des instruments. Suivant les termes de Hahn et Hausman (2003), même en utilisant les tests classiques de validité des instruments, « le chercheur peut estimer les « mauvais » résultats sans en avoir conscience » (p. 118). En outre, on notera aussi que l'estimateur VI peut être beaucoup moins efficace que le MCO.

4. Applications

Dans la présente section, nous décrivons deux applications de l'approche des effets de traitement pour l'évaluation ex post des interventions dans le domaine des transports. La première concerne une évaluation ex post des augmentations de la capacité routière urbaine aux États-Unis. L'objectif est d'estimer les ETM découlant de l'augmentation de la capacité routière en ce qui concerne la demande de trafic induite, les densités de trafic et la productivité. Cette étude utilise une méthodologie reposant sur le SP pour une estimation dose-réponse avec des traitements continus, proposée par Graham *et al.* (2014). La seconde application considère les impacts économiques régionaux de l'investissement dans le train à grande vitesse en Espagne, avec un estimateur DDD.

Évaluation *ex post* des augmentations de la capacité routière urbaine aux États-Unis

Objectif

Cette étude a pour objectif d'évaluer, au moyen des données longitudinales disponibles, les impacts des augmentations de capacité du réseau routier urbain sur les volumes de trafic, les densités de trafic et la productivité.

Méthode

L'étude utilise une méthodologie de régression à base de scores de propension généralisés (SPG) pour corriger la confusion et estimer les RPM et ETM. Les RPM considérés sont définis par $\mu(d) = E\,[Y_{it}(d)]$ et les ETM par $\tau\,(d) = E\,[Y_{it}(d)] - E\,[Y_{it}(0)]$, où d est une dose d'augmentation de capacité, i indexe les unités et t indexe le temps. On effectue les calculs pour un certain nombre de doses pertinentes et on construit une courbe dose-réponse au moyen d'une régression de spline pénalisée.

Comme on l'a vu ci-dessus, le SPG est défini par $\pi(D_{it} = d|X_{it}; \hat{a})$ et, pour la validité de l'inférence causale, nous exigeons l'indépendance conditionnelle et le support commun

$$Y_{it}(d) \perp D_{it} = d|X_{it} \quad \text{et} \quad Pr(D_{it} \in A|X_{it} = x_{it}) > 0 \; \forall \; x_{it}, A \subseteq C$$

où *C* est une région de support commun (voir par exemple Hirano et Imbens, 2004).

La convergence nécessite que X_{it} soit suffisant pour représenter la confusion, ce qui est toutefois, en réalité, une hypothèse intestable. Dans les applications longitudinales, on suppose souvent l'existence de co-variables inobservées, disons U_i et W_i, pouvant s'intégrer à notre modèle causal de telle sorte que :

$$D_{it} = f\,(X_{it}, U_i, W_i) \text{ et } Y_{it} \perp D_{it} = d|X_{it}, U_i.$$

Ainsi, U_i est une co-variable inobservée temporellement invariante confondante tandis que W_i est une co-variable inobservée temporellement invariante non confondante. Pour résoudre cette question, Graham *et al.* (2014) spécifient un modèle mixte longitudinal (ML) pour le SPG :

$$D_{it} = X_{it}^{\mathsf{T}}\theta_1 + b_i + H_{i,t-p}^{y\mathsf{T}}\theta_2 + \varepsilon_{it} \text{ avec } b_i \sim N(0, \sigma_\beta^2),$$

qui, en plus des confondeurs observés X_{it} variant dans le temps, conditionne sur les effets inobservés b_i au niveau des unités et un certain retard de la variable de réponse $H_{y\ i,t-p}$. Ainsi, le document propose pour l'estimation des ETM une approche à base de SPG qui prend un compte la confusion variant dans le temps mesurée, la confusion temporellement invariante inobservée et la bidirectionnalité entre réponse et traitement.

Les résultats analytiques et les simulations présentés dans le document montrent que, sous des conditions données, une approche SPG-ML fournira des estimations sans biais de la fonction dose-réponse, mais qu'un conditionnement plus étendu peut réduire l'efficacité et peut accroître la difficulté de trouver un chevauchement dans le support de la distribution des covariables.

L'algorithme pour l'estimation des ETM dans l'approche SPG-ML se déroule comme suit :

1. Estimer $f_{D|X,U}(D_{it}|X_{it}, b_i; \alpha)$ au moyen d'un modèle mixte ;
2. Utiliser â, avec une fonction de densité appropriée, pour calculer les SPG-ML pour les traitements $\pi(D_{it}|X_{it}, b_i; \hat{a})$ observés et $\pi(d|X_{it}, b_i; \hat{a})$ inobservés ;
3. Isoler une région de support commun : $\Pr(D_{it} \in A|X_{it}, b_i) > 0 \ \forall\ x_{it}, A \subseteq C$;
4. Estimer $E[Y_{it}|D_{it}, \pi(D_{it}|X_{it}, b_i \hat{a})]$ au moyen d'un modèle flexible ;
5. Faire la moyenne sur les valeurs prédites découlant de 4, évaluées à la dose d, pour obtenir une estimation ponctuelle du RPM à d : $\mu(d)$;
6. Répéter l'opération pour toutes les doses considérées, construire la courbe dose-réponse et estimer les ETM :

   $\tau\tau(d) = \mu(d) - \mu(0)$;

7. Appliquer un procédé de rééchantillonnage [*bootstrap* simple (par bloc)] sur les étapes 1 à 6 pour obtenir les erreurs types.

Données

Les données disponibles pour l'estimation sont tirées des chiffres de la mobilité urbaine du Texas Transportation Institute (TTI), qui décrivent la situation du trafic pour 101 grandes villes des États-Unis sur la période 1982-2007.

Figure 3. **Villes couvertes par les données sur la mobilité urbaine du Texas Transportation Institute**

Réponses : taux de variation annuel de la demande (kvp : kilomètre-véhicule parcouru), performance des réseaux (retard par kvp) et productivité (salaire moyen).

Traitement : taux de variation annuel du nombre de kilomètres de voies de circulation.

Covariables avant traitement (confondeurs) :

- **Réponses décalées** : pour saisir la causalité inverse ;
- **Congestion et volume de trafic** : mesurés par le retard et par kvp ;
- **Dimension et composition du réseau** : longueur du réseau, composition autoroutes / artères urbaines ;
- **Composition du trafic** : volume sur autoroutes / artères urbaines ;
- **Caractéristiques modales** : fréquentation des transports publics, prix des carburants par État ;
- **Économie** : productivité, revenu et structure économique ;
- **Emploi et répartition et croissance démographiques**.

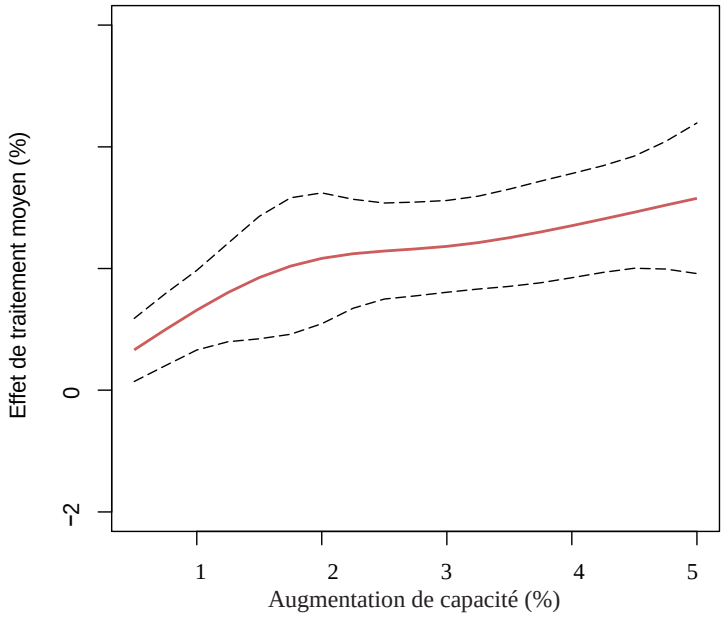
Confondeurs inobservés (inconnus) : caractéristiques à différents niveaux locaux ou régionaux, conception du réseau routier, comportement activité-déplacement.

Résultats

Les résultats pour nos trois réponses sont présentés dans les graphiques ci-dessous. Dans chacun d'eux, l'ETM est en ordonnée et la dose correspondante d'augmentation de capacité est en abscisse.

Examinons d'abord les résultats pour les volumes de trafic, mesurés en kvp.

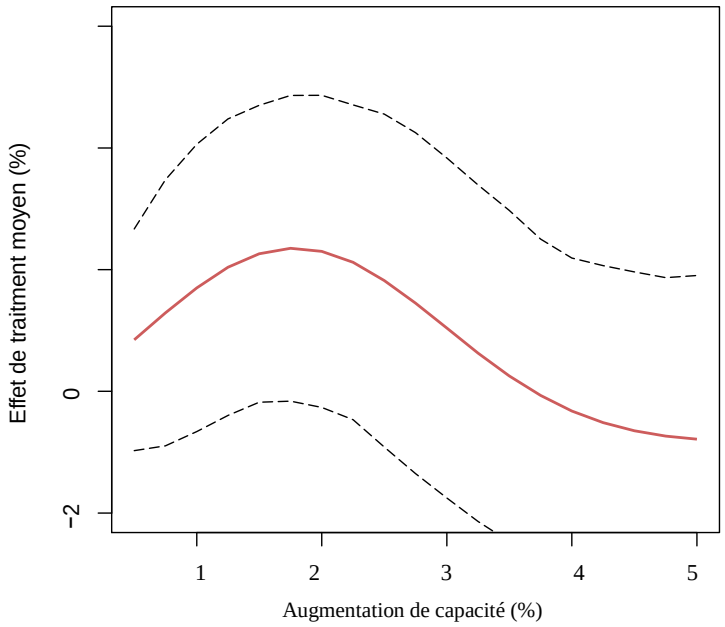
Figure 4. **Relation dose-réponse pour les volumes de trafic**



Les résultats indiquent l'existence d'une demande induite sur l'intervalle de doses, après correction de la confusion. L'ETM augmente plus rapidement que la capacité jusqu'à la dose de 2 % d'augmentation de capacité. En moyenne, on constate qu'une augmentation de 10 % du nombre de kilomètres de voies de circulation est associée à une augmentation de 9 % des kvp, nette de la « croissance naturelle » (estimée à 1.4 % par an). En conséquence, nous concluons que les augmentations de capacité dans l'intervalle considéré n'ont pas en général réduit la densité de trafic (c'est-à-dire le ratio entre le volume de trafic et la capacité).

Considérons ensuite les impacts sur la performance des réseaux, mesurée par le retard par kvp.

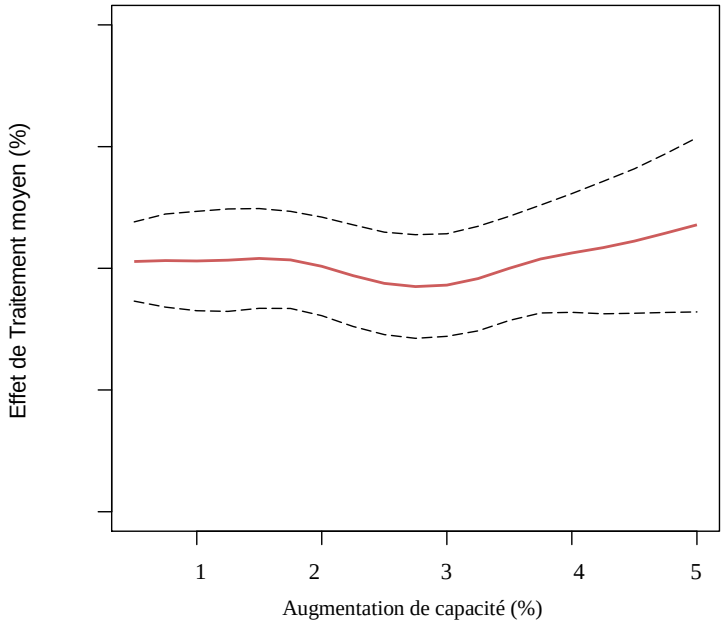
Figure 5. **Relation dose-réponse pour les performances du réseau (retard par kvp)**



Les résultats indiquent que les augmentations de capacité n'ont pas amélioré la congestion urbaine. L'usager de la route moyen n'a pas vu son retard changer à la suite des augmentations de capacité : on ne constate pas d'effets statistiquement significatifs sur le retard par kvp, et cela même pour les fortes augmentations de capacité. En fait, en raison de la croissance naturelle, la congestion s'est aggravée d'environ 3 % par an et, parce qu'il y a maintenant davantage de trafic, le total du retard urbain sur l'ensemble des doses augmente.

Enfin, regardons les effets sur la productivité, représentée par le taux de salaire urbain moyen.

Figure 6. **Relation dose-réponse pour la productivité (salaire moyen)**



Les résultats indiquent que les augmentations de capacité du réseau routier urbain n'ont pas induit une plus forte productivité. Si on effectue des régressions « naïves » de la productivité sur le traitement, on trouve bien une association positive entre la croissance de la capacité et les salaires, mais on ne constate pas d'ETM significatifs après ajustement pour la confusion et isolement d'une région de support commun.

Ainsi, il ressort de notre analyse causale que les augmentations de capacité du réseau routier urbain ont induit une demande mais n'ont pas amélioré la congestion ni élevé la productivité. Cela n'implique pas que les augmentations de capacité routière n'offrent pas en soi des avantages économiques. Ces résultats sont particuliers au cas de changements marginaux apportés à des réseaux urbains matures touchés par la congestion. Les augmentations de capacité ont permis une mobilité accrue, en ce sens qu'il y a plus de trafic, mais les coûts généralisés du réseau ne se sont pas améliorés et le retard urbain total a augmenté. L'effet d'échelle (trafic accru) ne semble pas avoir influé (en plus ou en moins) sur la productivité (soit +ve soit –ve).

Évaluation *ex post* des impacts économiques régionaux du train à grande vitesse en Espagne

Objectif

Entre 2000 et 2010, le gouvernement espagnol a réalisé le plus grand programme de construction de lignes à grande vitesse en Europe, de telle sorte qu'en 2011 le réseau espagnol à grande vitesse était devenu le plus grand d'Europe, dépassant la France et l'Allemagne. On prévoit que, d'ici 2020, 90 % de la population du pays habitera dans un rayon de 50 km d'une gare de ce réseau. On justifie souvent les investissements dans les projets de train à grande vitesse par les effets positifs escomptés pour la croissance économique régionale et nationale. Dans cette application, on effectue une analyse par DDD pour étudier les impacts sur la production économique découlant de la construction du

couloir ferroviaire à haute vitesse Madrid-Barcelone. C'est un travail de doctorat encore en cours, présenté ici pour illustrer l'évaluation ex post et non comme une conclusion définitive sur les impacts du train à grande vitesse (TGV) en Espagne.

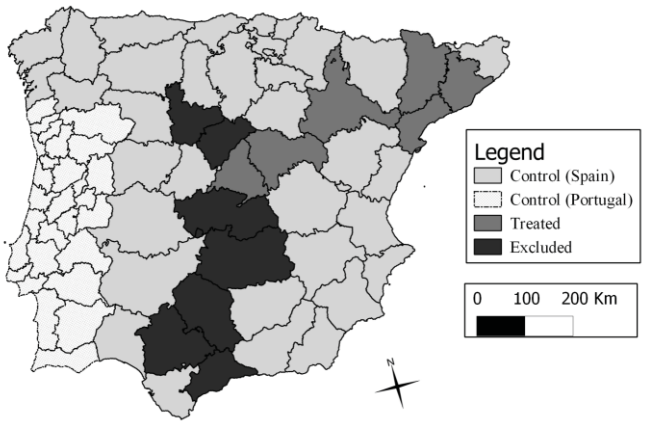
Méthode

Pour évaluer l'effet du couloir à grande vitesse reliant Madrid à Barcelone, nous divisons l'économie espagnole en 47 provinces continentales et nous considérons l'accès au TGV comme un traitement binaire. Les provinces recevant / ne recevant pas un accès au TGV sont dites traitées / non traitées et sont à la base de l'analyse par DDD. Elles sont indiquées dans le graphique 7. Nous utilisons la valeur ajoutée brute par habitant (VABph) comme variable de résultat dans la province i au temps t. Nous effectuons une régression de cette réponse sur la variable de l'année, une variable binaire pour chacun des groupes (est, ouest et nord) et les trois termes d'interaction entre l'année et les trois groupes témoins potentiels nord, est et ouest (année*est, année*ouest et année*nord) où le cas de base est le groupe des provinces traitées. Nous estimons aussi la même équation en ajoutant deux covariables pour tenir compte de la structure économique des différentes provinces : part des emplois dans le secteur manufacturier, l'énergie et la construction (part dans l'industrie) et part des emplois dans le secteur des services (part dans les services).

Résultats

Les résultats indiquent que les prédictions d'un impact positif sur les performances économiques des régions recevant le TGV ne se sont pas réalisées, au moins à court ou moyen terme. Dans le cas du couloir de TGV Madrid-Barcelone, l'analyse montre qu'il n'y a pas de différences significatives dans le profil de la croissance économique régionale avant et après le couloir le TGV entre les provinces traitées et non traitées.

Figure 7. **Provinces témoins et traitées pour le couloir nord-est de l'Espagne**



Légende Témoin (Espagne) Témoin (Portugal) Traitée Exclue

5. Conclusions

Dans ce document, nous avons examiné les méthodes visant à tirer des inférences causales des données observées et nous avons montré comment elles peuvent s'appliquer à l'évaluation ex post des projets dans les transports. Nous pensons qu'un modèle d'inférence causale fondé sur les résultats potentiels est très approprié à l'évaluation ex post parce qu'il est spécifiquement conçu pour des cas où les « traitements » ne sont pas assignés de manière aléatoire et où l'expérimentation n'est pas possible, circonstances qui caractérisent l'assignation des interventions dans les transports. Les méthodes examinées sont très utilisées pour l'analyse causale dans diverses disciplines scientifiques mais, à notre connaissance, elles ont fait l'objet de peu d'attention dans les analyses des transports. Nous présentons deux applications d'évaluation ex post reposant sur les techniques causales : l'une qui évalue les impacts des augmentations de capacité du réseau routier urbain aux États-Unis, et l'autre qui considère les impacts économiques régionaux des investissements dans le train à grande vitesse en Espagne.

Un avantage majeur des méthodes causales statistiques est qu'on peut les utiliser pour analyser l'impact des interventions sans poser de fortes hypothèses théoriques a priori concernant le comportement économique sous-jacent, comme l'exige l'ACA ex ante ou ex post. Toutefois, la validité de l'inférence causale à partir des données observationnelles nécessite elle-même un ensemble d'hypothèses assez strictes qui, dans de nombreux cas, peuvent n'être pas satisfaites par les données disponibles.

Bibliographie

Bang, H. et J. M. Robins (2005). Doubly robust estimation in missing data and causal inference models. Biometrics 61, 962–972.

Bertrand, M., E. Duflo et S. Mullainathan (2004). How much should we trust difference-in-differences estimates? Quarterly Journal of Economics 119 (1), 249–275.

Cox, D. R. (1958). Planning of experiments. Londres : John Wiley & Sons.

Fisher, R. A. (1935). The design of experiments. Édimbourg : Oliver and Boyd.

Graham, D.J., E.J. McCoy et D.A. Stephens (2014), Quantifying causal effects of road network capacity expansions on traffic volume and density via a mixed model propensity score estimator, *Journal of the American Statistical Association*, DOI:10.1080/01621459.2014.956871.

Hahn, J. et J. Hausman (2003), Weak instruments: diagnosis and cures in empirical economics, *American Economic Review*, 93, 118-125.

Hern´an, M.A. et J.M. Robins (2012), *Causal Inference. Monographs on Statistics and Applied Probability*, Londres : CRC Press.

Hirano, K. et G.W. Imbens (2004), The propensity score with continuous treatments, dans: Gelman et X.L. Meng (dir. pub.), *Applied Bayesian modeling and causal inference from incomplete data perspectives*, pp. 73-84. New York : Wiley.

Holland, P.W. (1986), Statistics and causal inference, *Journal of the American Statistical Association*, 81 (396), 945-970.

Imbens, G.W. (2000), The role of the propensity score in estimating dose-response functions, *Biometrika*, 87 (3), 706-710.

Imbens, G.W. et J.M. Wooldridge (2009), Recent developments in the econometrics of program evaluation, *Journal of Economic Literature*, 47 (1), 5-86.

Joffe, M.M. et P.R. Rosenbaum (1999), Propensity scores, *American Journal of Epidemiology*, 150 (4), 327-333.

Kang, J.D.Y. et J.L. Schafer (2007), Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data, *Statistical Science*, 22 (4), 523-539.

Lunceford, J.K. et M. Davidian (2004), Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study, *Statistics in Medicine*, 23, 2937-2960.

Neyman, J. (1923), On the application of probability theory to agricultural experiments. Essay on principles, Section 9, *Statistical Science*, 5 (4), 465-480. Traduit en 1990.

Robins, J.M. (2000), Robust estimation in sequentially ignorable missing data and causal inference models, dans : *Proceedings of the American Statistical Association*, Section on Bayesian Statistical Science, pp. 6-10, Alexandria, VA : American Statistical Association.

Robins, J.M. et A. Rotnitzky (2001), Comment on "Inference for semiparametric models: some questions and an answer", *Statistical Sinica*, 11, 920-936.

Robins, J.M., A. Rotnitzky et M.J. van der Laan (2000), Comment on the Murphy and Van der Vaart article "On profile likelihood", *Journal of the American Statistical Association*, 95, 431-435.

Rosenbaum, P.R. et D.B. Rubin (1983), The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70 (1), 41-55.

Rubin, D.B. (1973a), Matching to remove bias in observational studies, *Biometrics*, 29, 159-183.

Rubin, D.B. (1973b), The use of matched sampling and regression adjustments to remove bias in observational studies, *Biometrics*, 29, 185-203.

Rubin, D.B. (1974), Estimating causal effects of treatments in randomized and non-randomized studies, *Journal of Educational Psychology*, 66 (5), 688-701.

Rubin, D.B. (1977), Assignment to treatment group on the basis of a covariate, *Journal of Educational Statistics*, 2 (1), 1-26.

Rubin, D.B. (1978), Bayesian inference for causal effects: the role of randomization, *Annals of Statistics*, 6 (1), 34-58.

Rubin, D.B. (1980), Comment on 'Randomization analysis of experimental data in the Fisher randomization test' by Basu, *Journal of the American Statistical Association*, 75 (371), 591-593.

Rubin, D.B. (1986), Comment: which ifs have causal answers? *Journal of the American Statistical Association*, 81 (396), 961-962.

Rubin, D.B. (1990), Neyman (1923) and causal inference in experiments and observational studies, *Statistical Science*, 5 (4), 472-480.

Tsiatis, A.A. (2006), *Semiparametric theory and missing data*, Berlin : Springer.

Tsiatis, A.A. et M. Davidian (2007), Comment: Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data, *Statistical Science*, 22 (4), 569-573.

van der Laan, M. et J.M. Robins (2003), *Unified methods for censored longitudinal data and causality*, Berlin : Springer.

Forum International des Transports

Forum International des Transports
2 rue André Pascal
75775 Paris Cedex 16
itf.contact@oecd.org
www.internationaltransportforum.org