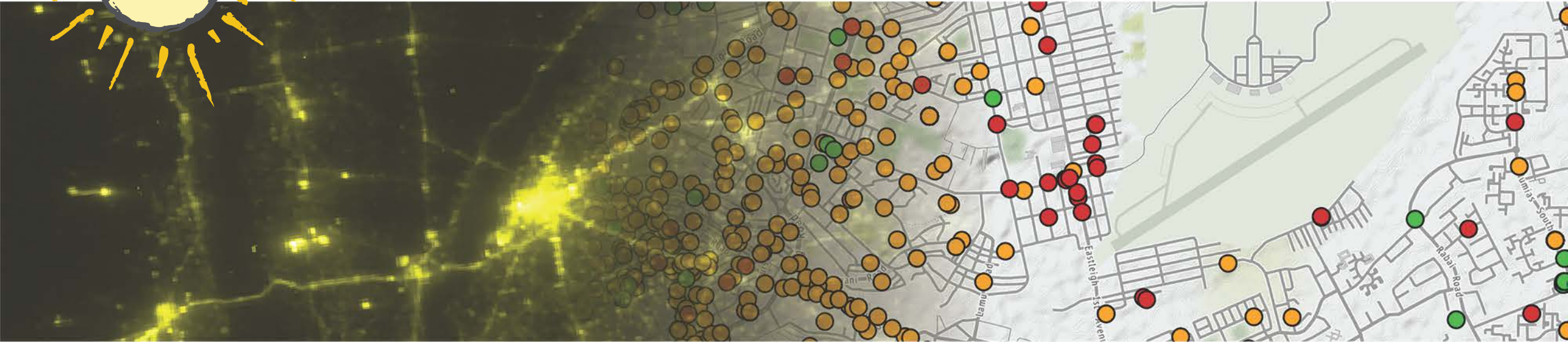




IE CONNECT FOR IMPACT

Transforming the Growth Potential
of Transport Investments

Mesurer l'impact I: Inférence Causale et Méthodes Quasi-Expérimentales

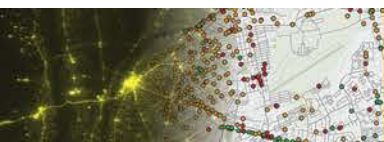


Introduction

- Objectif principal:
 - Guider les décisions politiques et de programmes sur base des preuves solides.
- Argument:
 - Plusieurs types d'évidence du plus simple & moins-couteux au plus complexe.
 - En cas des doutes sur l'efficacité d'un programme, l'étude d'incidence (ou évaluation d'impact) se justifie.
 - Si nous n'avons pas des doutes nous pouvons nous passer de l'étude d'incidence.

Pourquoi l'évaluation d'impact?

- Permet de *mesurer les* impacts directs des politiques publiques, programmes d'infrastructures...
 - Est-ce que le programme à un impact sur l'indicateur choisi ?
 - Comprendre les avantages nets du programme & comparer les différentes options du programme
 - Comprendre la répartition des effets du programme entre différent sous-groupes
 - Comprendre comment les résultats peuvent changer en modifiant certains aspects d'un programme
- Autres avantages opérationnels
 - Décaissement plus rapide & moins de retards dans la mise en œuvre du projet

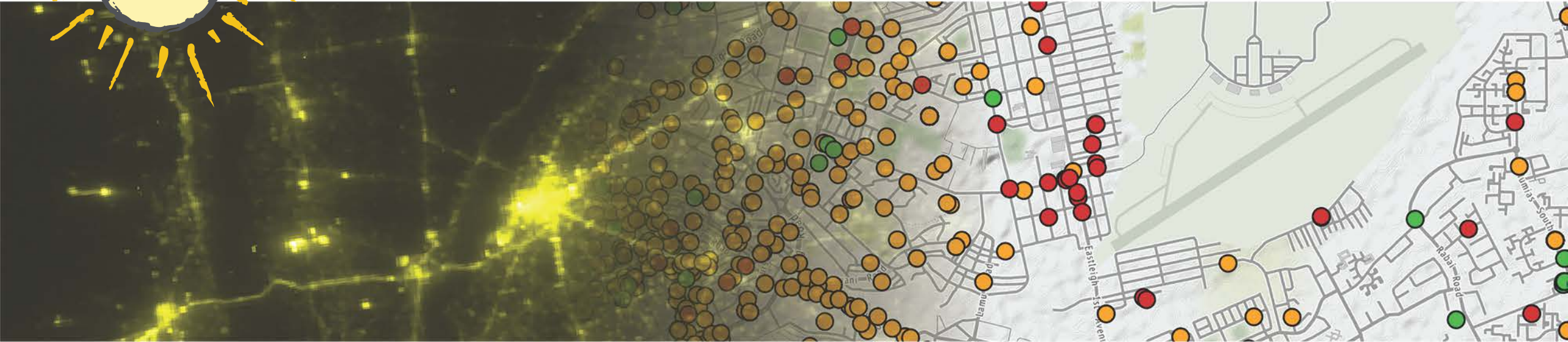




IE CONNECT FOR IMPACT

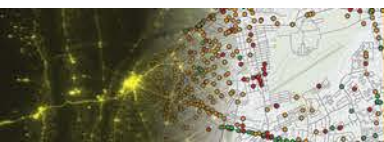
Transforming the Growth Potential
of Transport Investments

Évaluation d'impact et inférence



Inférence causale?

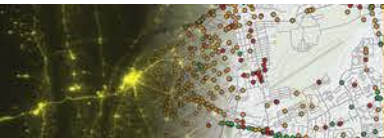
- L'inférence causale est la pierre angulaire des évaluations d'impact
- Les politiques de développement invoquent généralement des questions/rerelations *de cause à effet*.



Exemple: Les programmes d'emploi pour les jeunes peuvent-ils réduire le risque de participation à la violence?



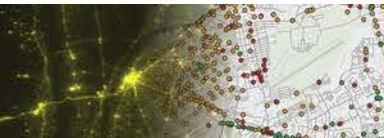
Exemple: Les programmes de formation des agents de l'État entraînent-ils une amélioration des performances?



Exemple: Les garderies peuvent-elles accroître la participation des femmes dans la vie économique?



Exemple: La modernisation des routes augmente-t-elle le bien être de la population rurale?





IE CONNECT FOR IMPACT

Transforming the Growth Potential
of Transport Investments

Deux approches

1. Suivi et évaluation (traditionnel)

2. Évaluation d'impact



Suivi et évaluation (traditionnel)

- Mesurer les progrès des indicateurs au fil du temps:
 - « *Est-ce que le projet est sur la bonne voie ?* »
- L'accent est mis sur les activités du projet et les bénéficiaires
- Permet d'observer le changement, mais ne nous dit pas pourquoi ou comment il est arrivé.



Le défi de l'attribution de l'impact II: facteurs externes

Les impacts du programmes se confondent avec des effets locaux, nationaux et globaux



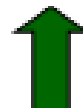
IMPACT



RESULTATS



EXTRANTS



INTRANTS

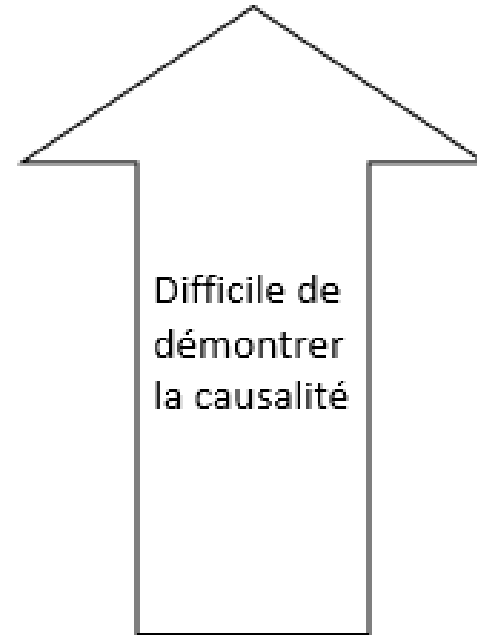
Le services atteints les bénéficiaires



Paramètres sous contrôle du programme



Difficile de démontrer la causalité



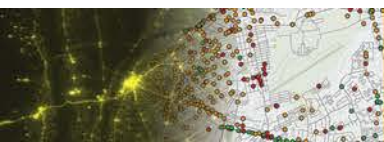
Évaluation d'impact

- Impact ?
- « *Qu'arriverait-il en l'absence du programme ou projet ?* »
- C'est la différence entre les résultats avec programme et sans programme
- L'objectif de l'évaluation d'impact est de mesurer cette différence de manière à *l'attribuer au programme*, et *seulement au programme*



Ce dont on a besoin

- **Idéalement** : les résultats **avec** et **sans** programme pour la même unité d'analyse, qu'elle soit un individu, un village, un groupe ...
- **Problème:**
 - On ne peut pas à la fois recevoir et ne pas recevoir le traitement
 - Il y a un problème de **manque de données**
- **Solution:**
 - Avoir un **groupe de comparaison**/témoin pour estimer ce qui serait arrivé sans programme
 - Ne peut être observé et doit donc être constitué



Solution: trouver un contrefactuel

- Trouver un groupe de comparaison, le **contrefactuel**:

Ce qui s'est vraiment passé



Monde réel !!!

Ce qui se serait passé sans programme - contrefactuel



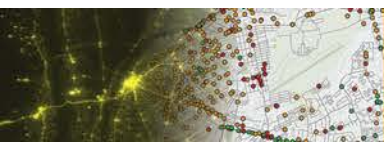
Monde imaginaire !!!

La clé d'une bonne évaluation d'impact est un scénario contrefactuel valide !

Comment trouver un contrefactuel valide?

Caractéristiques d'un contrefactuel valide

- Le groupe de traitement et le groupe témoin doivent avoir les mêmes caractéristiques
 - observables
 - non observables
- Donc le changement dans les résultats ou l'indicateur peut être attribué seulement au programme



L'expérience parfaite

- Quel est l'impact de donner à Babouche de l'argent de poche supplémentaire sur le nombre de bottines rouge qu'il possède ?
- Imaginons que Babouche ait un clone:
 - Identique à l'extérieur (observables)
 - Identique à l'intérieur (non observables)

Babouche



Clone



Nous sommes tous les deux des singes

Nous aimons tous les deux les bottines rouges

L'expérience parfaite

- Si on applique le traitement à Babouche – si on lui donne plus d'argent de poche

Babouche



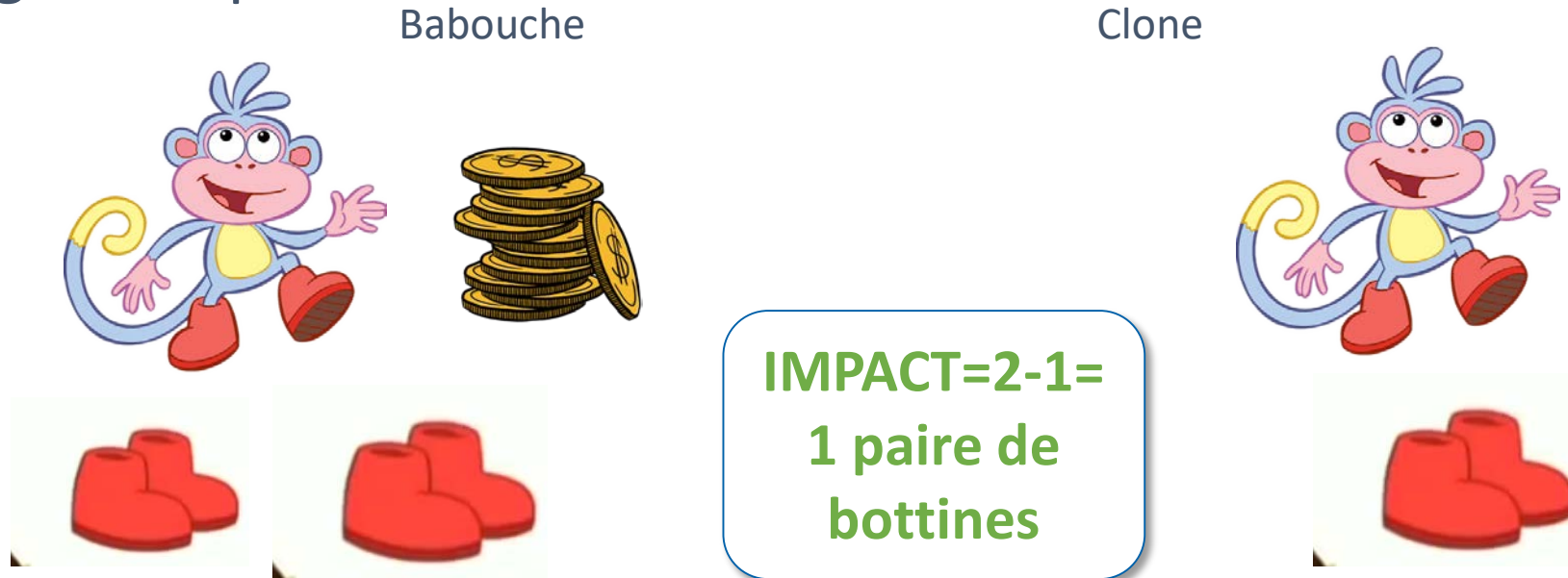
Clone



La seule différence est le projet ou traitement

L'expérience parfaite

- Si on applique le traitement à Babouche – si on lui donne plus d'argent de poche



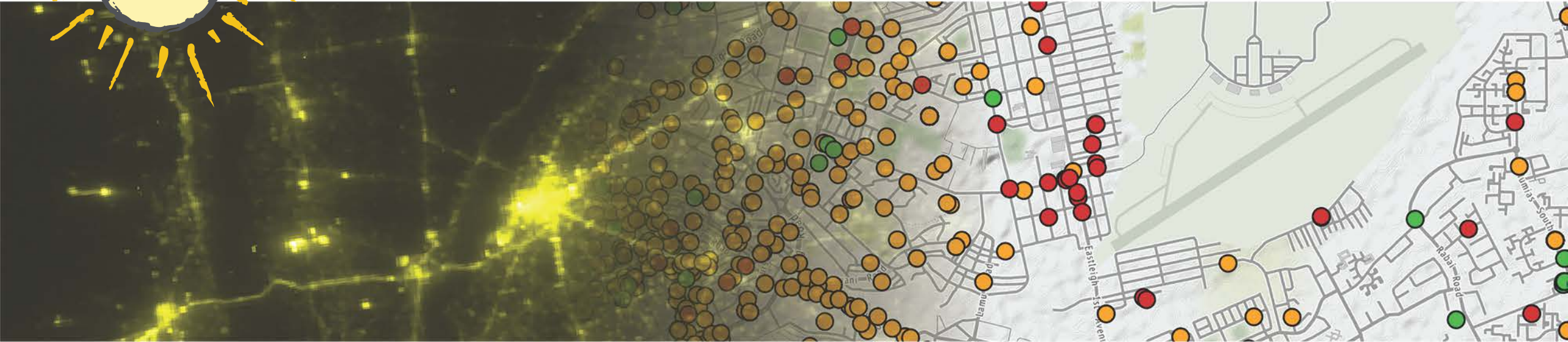
*Comme Babouche et son clone sont identiques,
la différence du nombre de paires de bottines est
due au traitement*



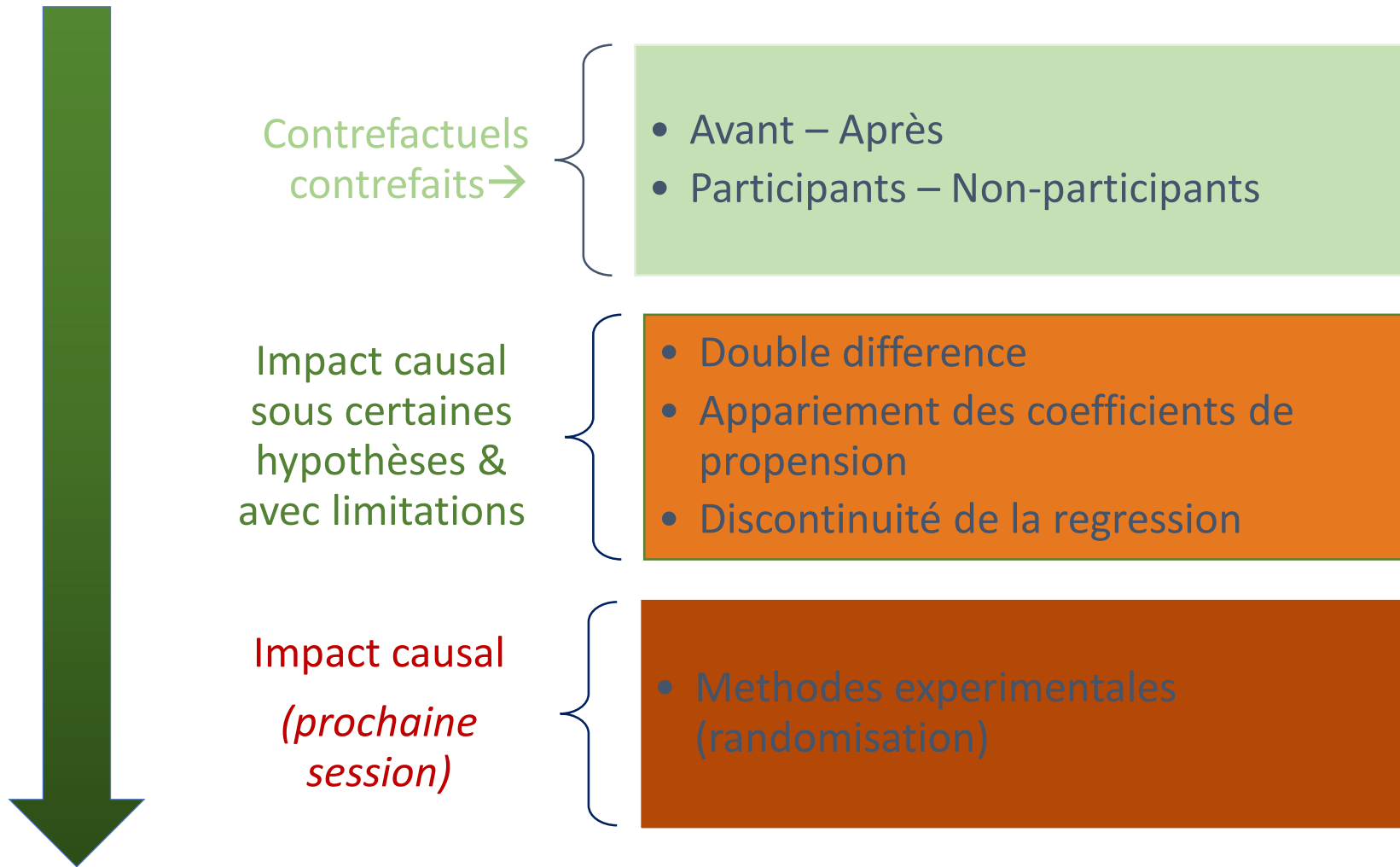
IE CONNECT FOR IMPACT

Transforming the Growth Potential
of Transport Investments

Évaluation d'impact en pratique



Comment construire un contrefactuel valide pour mesurer l'impact causal ?



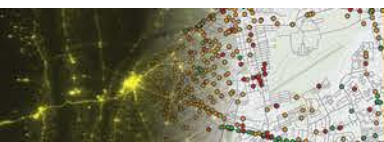
Étude de cas: modernisation des routes rurales et bien être

- **Problème:** La République d'Atlantis veut moderniser les routes Rurales pour améliorer le bien-être socio- économique des citoyens
- **Intervention:** Le gouvernement décide de lancer un programme pilote pour voir si ce programme sera efficace
- **Résultats clés:** Bien-être socio-économique, plus particulièrement la consommation par habitants des ménages ruraux.
- **Déroulement :**
 - **Janvier 2014:** Parmi les 9000 villages ruraux, 2,000 sont invités à participer
 - **Mars 2014- mars 2015:** 1021 villages déposent leur dossier à temps et voient leur route être remise à niveau . Les 979 restants sont les non participants.
 - **Juin 2016:** Enquête auprès des ménages des 2000 villages invités

Contrefactuels contrefaits I:

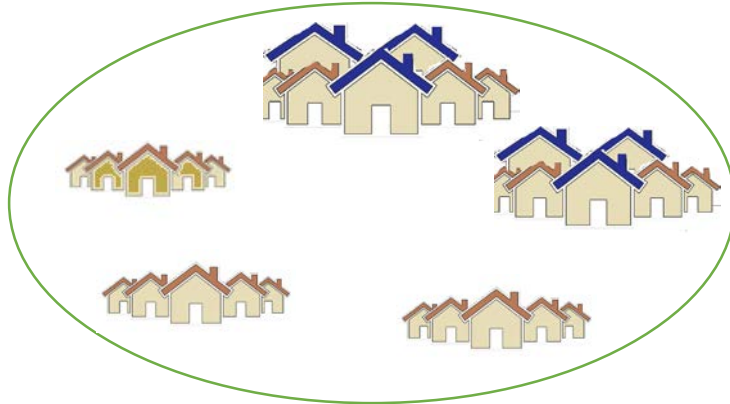
Comparaison participants et non-participants

- **Idée:**
 - **Comparer** la consommation des ménages des villages participants **avant et après** modernisation des routes
- **Hypothèse fondamentale:**
 - Si le programme n'avait pas existé, le résultat pour les participants aurait suivi la même tendance que dans le passé

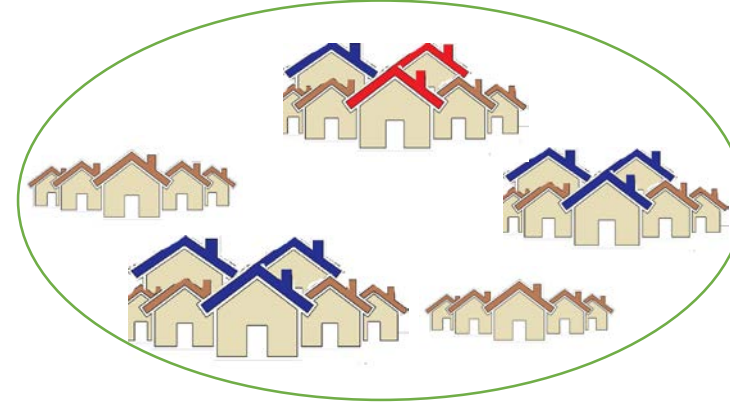


Contrefactuels contrefaits I: Comparaison avant-après

Participants 2014



Participants 2016



	Participants 2014	Participants 2016	Différence
Consommation dans les villages participants	274.4	301.6	27.2 *

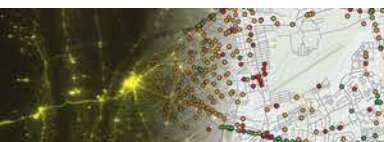
La différence represent-elle une estimation non-biaisée de l'impact programme?

Pas forcément: différence dans le temps

- Pas que le programme: trop de facteurs changent avec le temps, par ex. amélioration de la conjoncture économique,...
- Des changements surviennent avec le temps, même sans programme.

Contrefactuels contrefaits I: Comparaison participants et non-participants

- **Idée:**
 - **Comparer** la consommation des ménages des villages **avec** routes modernisées **et** des villages **sans**
- **Hypothèse fondamentale:**
 - Si le programme n'avait pas existé, le résultat pour les participants serait le même que pour les non-participants



Comparaison participants et non-participants



	Participants	Comparaison	Différence
Consommation des ménages	301.6	219.1	82.5*

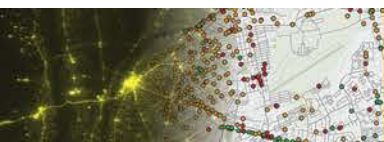
La différence représente-elle une estimation non-biaisée de l'impact du programme?

Pas forcément: biais de selection

- Les villages participants sont plus denses, ont des plus hauts niveau de consommation (observable)
- Villages ont une organisation différente, ont des histoires différentes (non observables)

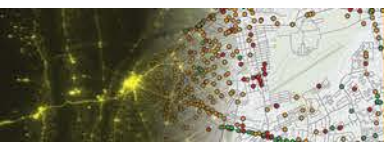
Méthode d'appariement

- **Idée:** Pour chaque unité traitée, choisir la meilleure unité de comparaison (*match*) sur la base des similarités dans les caractéristiques observées
- **Contrefactuel:** groupe des non-participants avec les mêmes caractéristiques observables que les participants
- **Hypothèse fondamentale: après appariement...**
 - Pas de différence systématique dans les caractéristiques non observées influençant la participation et le résultat
 - les caractéristiques non observables n'affectent pas l'assignation au traitement ni les résultats étudiés



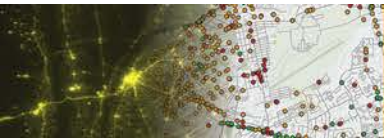
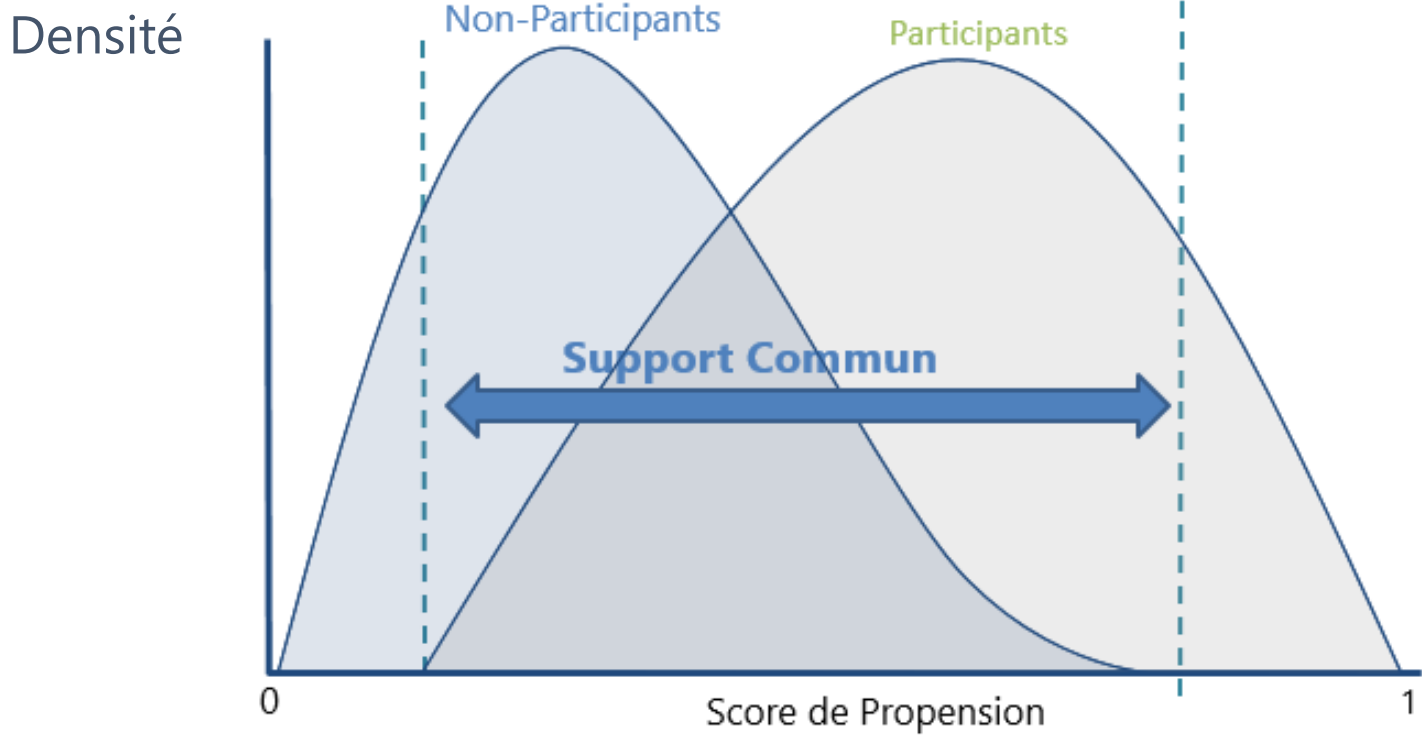
Méthode d'appariement

- **Apparier sur la base du «score des coefficients de propension»:**
 - Calculer la probabilité de participation de chacun, basée sur leurs caractéristiques observables.
 - Pour chaque participant, trouver un échantillon de non participants avec un score de propension similaire.
 - Comparer les indicateurs des résultats pour chaque observation et le groupe d'observations appariées.



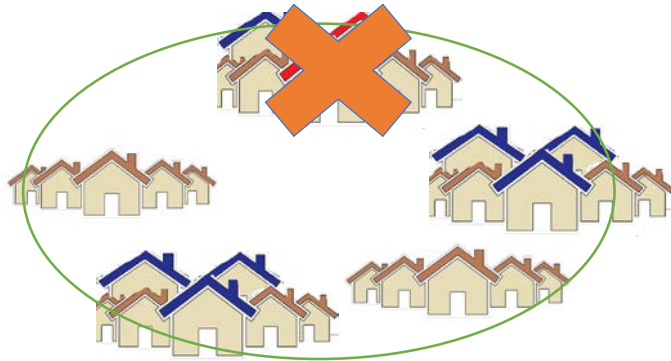
Méthode d'appariement...

Densité du score de propension:

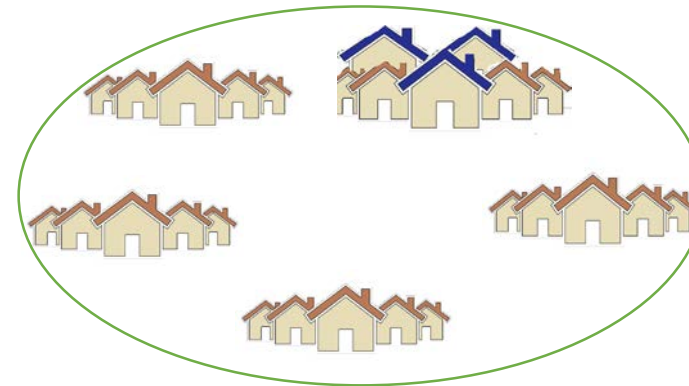


Methode d'appariement...

Participants



Non-participants



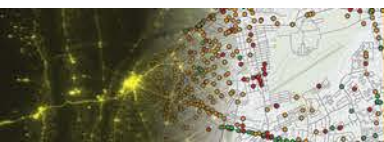
	Participants	Comparaison	Différence
Consommation des ménages	290.23	234.41	55.8*
Nombre d'observations utilisées	886	751	

- La différence représente-elle une estimation *non-biaisée* de l'impact du programme?
- **Pas forcément: Influence des facteurs non-observables**

Méthode de double différence (DD)

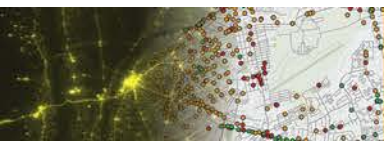
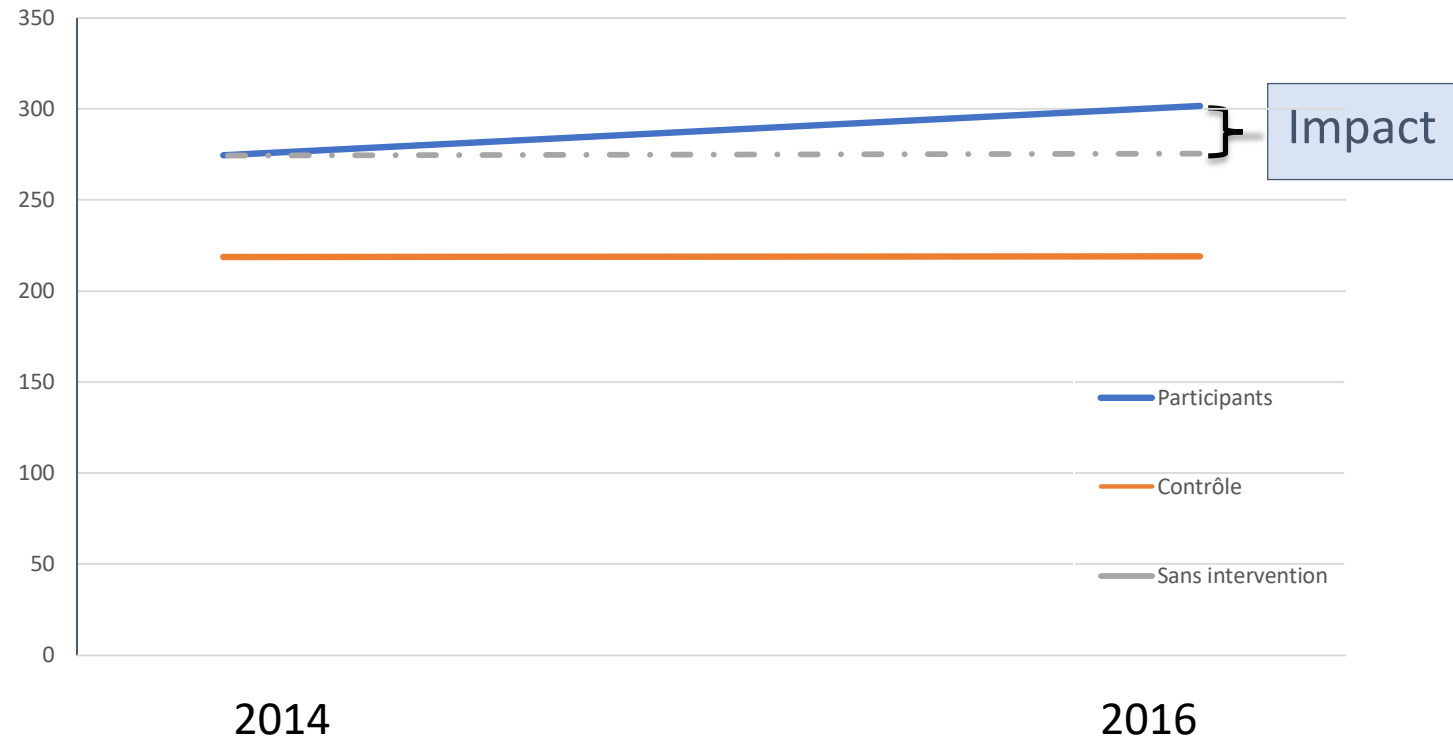
La méthode DD:

- Compare les différences de résultats entre participants et non-participants dans le programme au fil du temps
- **Hypothèse d'identification:**
 - “Tendances parallèles” en l'absence du programme
- **Le contrefactuel**
 - *C'est le changement au fil du temps* pour les non-participants dans le programme
- (Sous certaines hypothèses), la DD peut produire des estimations moins biaisées



Hypothèse: tendances parallèles

Consommation des ménages



Example : DD et routes rurales

	Traitement	Comparaison	Différence
POST- modernisation Consommation en 2016	301.6	219.1	82.5
PRÉ-modernisation Consommation en 2014	274.4	219	55.4
Différence de niveau de consommation en 2016 et 2014	27.2 (301.6-274.4)	0.1 (219.1-219)	27.1* (301.6-274.4)-(219.1-219) =(Différence-dans la - Différence)

Méthode de Régression Discontinue (RD)

Beaucoup de programmes sociaux sélectionnent les bénéficiaires en utilisant un indice ou un score:

Programmes de lutte
contre la pauvreté



Ciblent les ménages sous un seuil de
pauvreté

Retraites



Ciblent la population au dessus d'un
certain âge

Bourses d'étude



Destinées aux élèves dont les résultats
aux test sont élevés

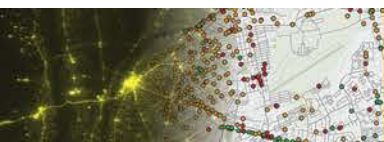
Agriculture



Fertilisant destiné à des petites
exploitations (< un certain nombre
d'hectares)

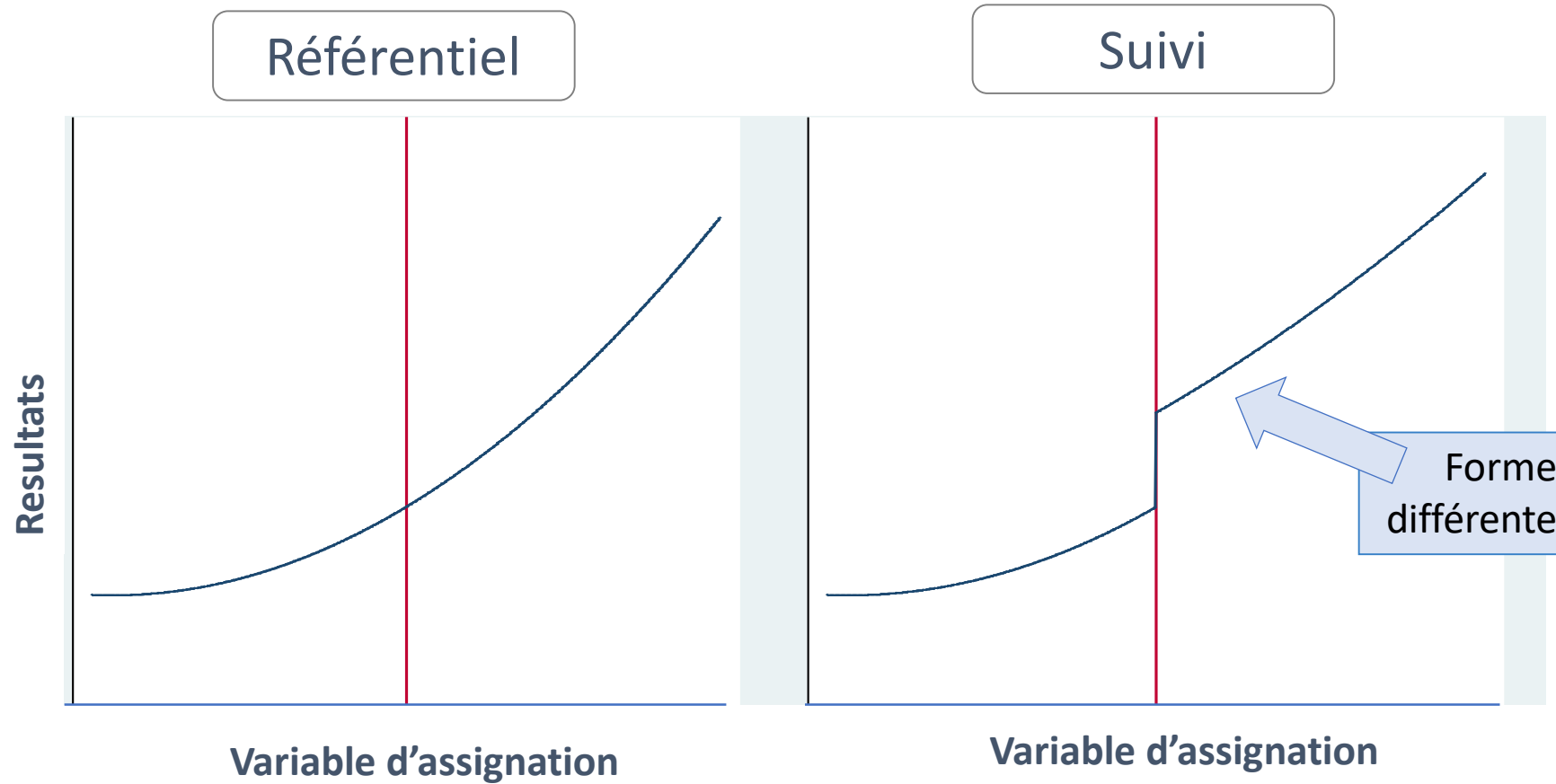
Méthode de Régression Discontinue (RD)...

- **Hypothèse fondamentale:**
 - Les unités juste au-dessus du seuil sont comparables à celles juste au-dessous
- **Le RD se base sur la compréhension du processus de sélection:**
 - Etablir une règle de sélection claire & un score quantifiable simple et continu
 - L'assignation du programme est basée sur base d'un seuil
 - Cibler les unités autour du seuil pour l'évaluation



Méthode de Regression Discontinue (RD)

Nous sommes à la recherche de ce type de tendances...



Un exemple de RD: Effet de programme de bourses rurales

- **Objectif**

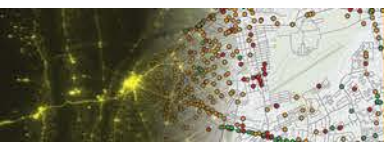
- Améliorer la diversification des revenus parmi les femmes des ménages ruraux les plus pauvres

- **Méthode**

- Ménages avec un score (indice de richesse) ≤ 50 sont pauvres
- Ménages avec un score (indice de richesse) > 50 sont non pauvres

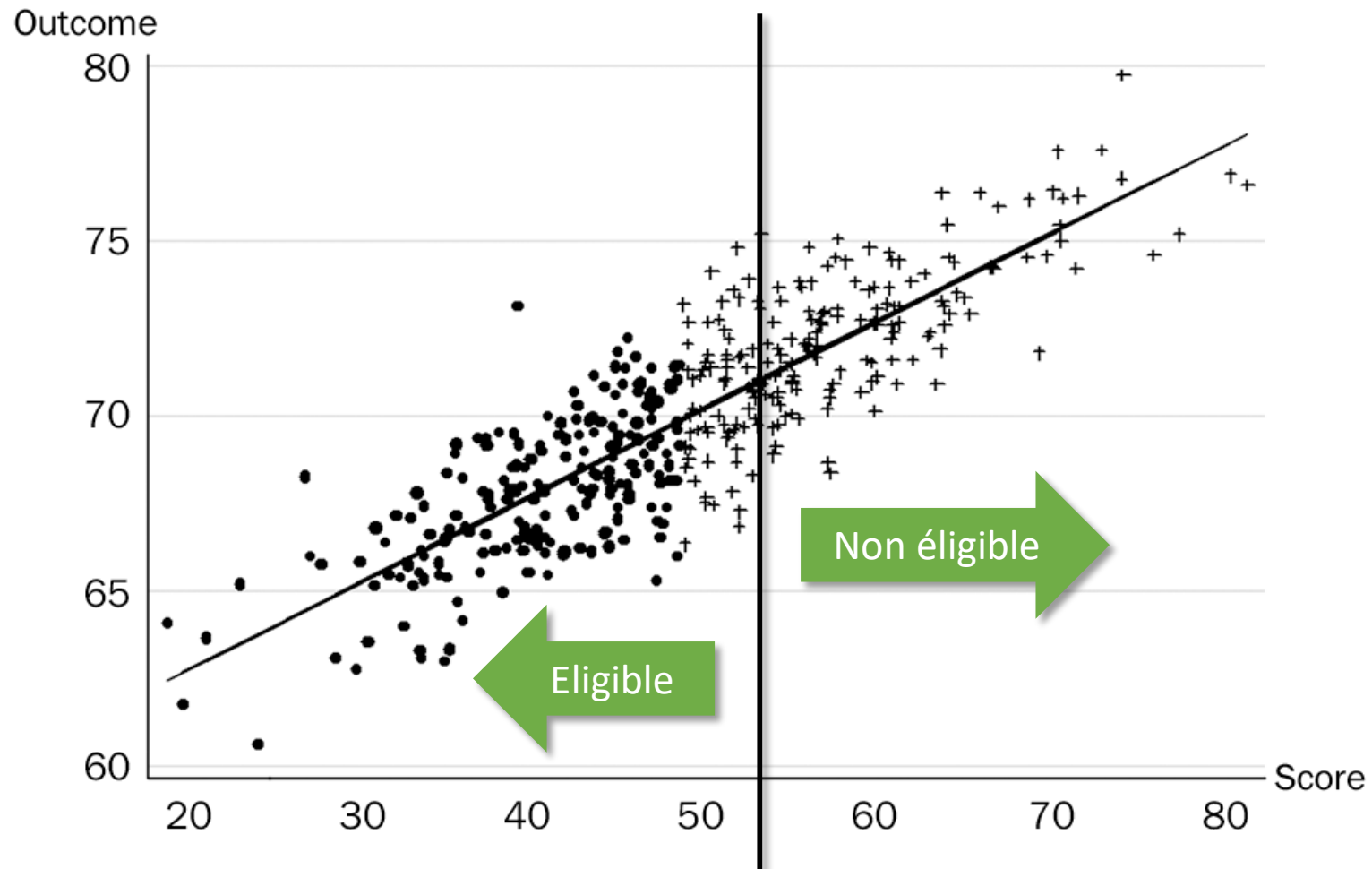
- **Intervention**

- Les femmes faisant partie des ménages pauvres reçoivent une bourse pour le transport et la commercialisation de leurs produits.



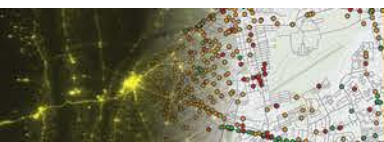
Méthode de Regression Discontinue (RD)...

Situation de référence:



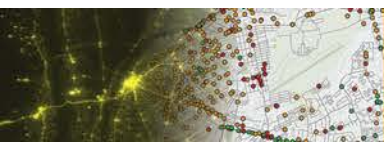
Synthèse

- RD se prête à l'évaluation prospective lorsque la randomisation n'est pas faisable:
 - Stratégie applicable à tout programme qui se base sur un **critère d'éligibilité**
 - Possibilité d'exploiter plusieurs seuils pour améliorer la validité externe
 - **L'effet est causal** mais **local** et donc il y a un **problème de généralisation**.



Synthese: méthodes non-expérimentales

- Toutes les situations ne se prêtent pas à la même méthode
- Différentes méthodes produisent des résultats différents
- Les méthodes faibles peuvent conduire à des résultats biaisés
- Certaines méthodes non-expérimentales nécessitent de hypothèses fortes et de très bonnes données
- Les résultats de l'évaluation d'impact sont valides seulement si nous utilisons des méthodes rigoureuses.



Merci !

