

# Quels accès aux données individuelles ? L'enjeu pour l'évaluation des politiques publiques

Antoine Bozio

*Institut des politiques publiques (IPP)*

PSE-École d'économie de Paris  
Paris – 31 mai 2018

- I. L'importance des données individuelles pour l'évaluation
- II. L'accès aux données administratives
- III. Quels progrès restent à faire ?

# I. L'importance des données individuelles pour l'évaluation

- **Objectif de l'évaluation : approfondir nos connaissances**
  - Mesurer les impacts des politiques publiques
  - Comprendre les mécanismes en jeu
  - Réfléchir à l'architecture optimale

# I. L'importance des données individuelles pour l'évaluation

- **Objectif de l'évaluation : approfondir nos connaissances**
  - Mesurer les impacts des politiques publiques
  - Comprendre les mécanismes en jeu
  - Réfléchir à l'architecture optimale
- **Le problème de l'évaluation d'impact**
  - But : approcher la causalité des politiques publiques
  - Construire un contrefactuel convaincant

# I. L'importance des données individuelles pour l'évaluation

- **Objectif de l'évaluation : approfondir nos connaissances**
  - Mesurer les impacts des politiques publiques
  - Comprendre les mécanismes en jeu
  - Réfléchir à l'architecture optimale
- **Le problème de l'évaluation d'impact**
  - But : approcher la causalité des politiques publiques
  - Construire un contrefactuel convaincant
- **La nécessité des données individuelles**
  - Identifier des individus (ou entreprises) similaires mais non touchés
  - Besoin d'échantillons de données individuelles

## ① Les données d'enquêtes

- Données de la Statistique publique
- Indispensables pour décrire de façon fine l'état du pays
- Des échantillons limités par construction

## ① Les données d'enquêtes

- Données de la Statistique publique
- Indispensables pour décrire de façon fine l'état du pays
- Des échantillons limités par construction

## ② Les données “administratives”

- Les données d'origine privée
  - Données de scanner (prix, consommation, etc.) ;
  - Données de plateformes internet (loyers, prix immobilier, etc.)

## ① Les données d'enquêtes

- Données de la Statistique publique
- Indispensables pour décrire de façon fine l'état du pays
- Des échantillons limités par construction

## ② Les données “administratives”

- Les données d'origine privée
  - Données de scanner (prix, consommation, etc.) ;
  - Données de plateformes internet (loyers, prix immobilier, etc.)
- Les données d'origine publique
  - Données de santé (consommation de soin, etc.)
  - Données fiscales
  - Données de sécurité sociale



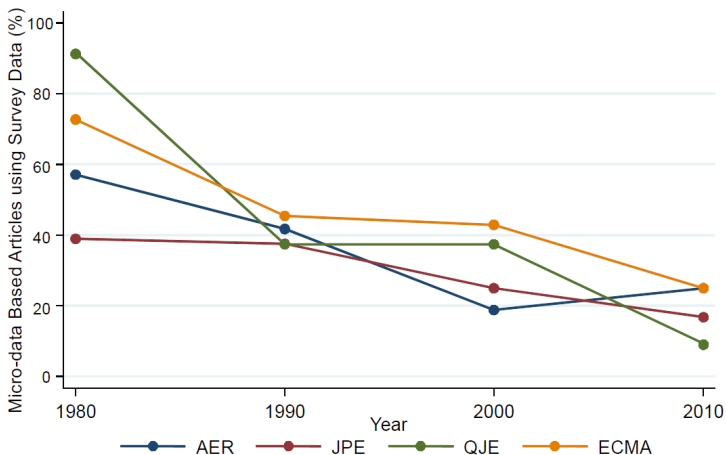
## ① Les données d'enquêtes

- Données de la Statistique publique
- Indispensables pour décrire de façon fine l'état du pays
- Des échantillons limités par construction

## ② Les données “administratives”

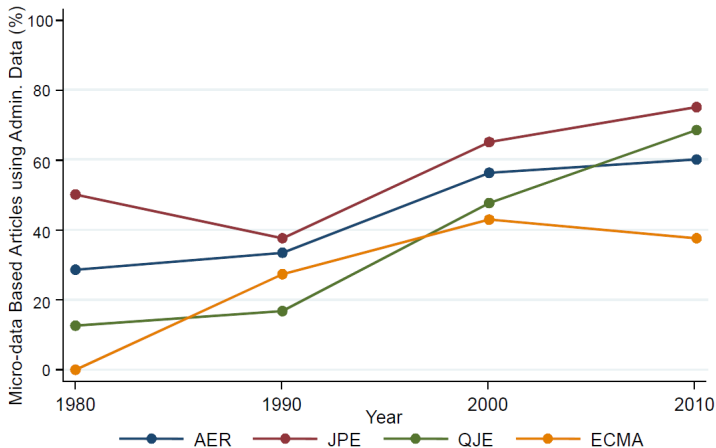
- Les données d'origine privée
  - Données de scanner (prix, consommation, etc.) ;
  - Données de plateformes internet (loyers, prix immobilier, etc.)
- Les données d'origine publique
  - Données de santé (consommation de soin, etc.)
  - Données fiscales
  - Données de sécurité sociale
- Développement de méthodes d'évaluation d'impact exploitant ces “big data”

**FIGURE 1:** Utilisation de données d'enquête dans les articles publiés, 1980–2010



SOURCE : Chetty and Bruich (2012), Public Economics Lectures.

**FIGURE 2:** Utilisation de données administratives dans les articles publiés, 1980–2010



SOURCE : Chetty and Bruich (2012), Public Economics Lectures.

## II. Accès aux données individuelles

- **Le retard français**

- Avance très nette des pays nordiques (Suède, Danemark, Norvège)
- Production de données administratives de qualité en France
- Problème d'accès : données protégées par des secrets

- **Le retard français**
  - Avance très nette des pays nordiques (Suède, Danemark, Norvège)
  - Production de données administratives de qualité en France
  - Problème d'accès : données protégées par des secrets
- **Les avancées législatives récentes**
  - Les données statistiques (juillet 2008)
  - Les données fiscales (loi du 22 juillet 2013)
  - Les données de santé (loi du 26 janvier 2016)
  - Loi pour une République numérique (oct. 2016)

- **Le retard français**

- Avance très nette des pays nordiques (Suède, Danemark, Norvège)
- Production de données administratives de qualité en France
- Problème d'accès : données protégées par des secrets

- **Les avancées législatives récentes**

- Les données statistiques (juillet 2008)
- Les données fiscales (loi du 22 juillet 2013)
- Les données de santé (loi du 26 janvier 2016)
- Loi pour une République numérique (oct. 2016)

- **Références**

- *Note IPP n°25, "Évaluation des politiques publiques : le bilan contrasté du quinquennat" (Mars 2017)*
- *Bozio et Geoffard (2017), L'accès des chercheurs aux données administratives – État des lieux et propositions d'actions, Rapport du groupe de travail du Cnis.*

## ① Des dispositifs législatifs complexes et mal connus

- Diffuser un vade-mecum de l'état du droit à destination des administrations et des chercheurs
- Rapprocher les procédures législatives d'accès et de traitement des données couvertes par un secret
- Inciter les administrations à saisir le comité du secret

## ① Des dispositifs législatifs complexes et mal connus

- Diffuser un vade-mecum de l'état du droit à destination des administrations et des chercheurs
- Rapprocher les procédures législatives d'accès et de traitement des données couvertes par un secret
- Inciter les administrations à saisir le comité du secret

## ② Améliorer les modalités pratiques d'accès aux données

- Mettre en place une instance de concertation unique
- Engager un processus de recensement des données existantes
- Prévoir des procédures d'accès sécurisés qui permettent l'appariement des données



## ③ Développer un modèle économique pérenne de l'accès aux données

- Prévoir des budgets dans les COG pour les travaux d'extraction des données de gestion
- Encourager des exploitations exploratoires de données brutes
- Les financements publics de la recherche doivent prendre en compte les coûts d'accès aux données

## ③ Développer un modèle économique pérenne de l'accès aux données

- Prévoir des budgets dans les COG pour les travaux d'extraction des données de gestion
- Encourager des exploitations exploratoires de données brutes
- Les financements publics de la recherche doivent prendre en compte les coûts d'accès aux données

## ④ Améliorer les pratiques d'archivage

- Rappeler aux administrations que les délais de conservation des données fixés par la Cnil n'imposent pas la destruction des données, mais l'archivage de ces données

# Quels accès aux données individuelles ? L'enjeu pour l'évaluation des politiques publiques

Antoine Bozio

*Institut des politiques publiques (IPP)*

PSE-École d'économie de Paris  
Paris – 31 mai 2018