

# **Les Besoins des démo-économistes en matière d'appariements sécurisés**

Nathalie Picard (THEMA, Université de Cergy-  
Pontoise, INED et Ecole Polytechnique)

Benoît Riandey (INED), Anne Solaz (INED)



# Un contexte plus encourageant qu'on ne pourrait le craindre

« Appariements sécurisés :

Statisticiens, ayez de l'audace ! »

S.V-T, 13 novembre 2006

# Données, besoins et solutions

1. Les grands panels de la statistique publique
  - Les Bases de données médico-économiques
  - Appariement des données sur l'emploi ?
  - Engager une réflexion collective
  - Nouvelle architecture des appariements
2. Articuler les données sur les ménages
  - Les données souhaitées
  - Recherchons grand panel de ménages ?
  - Micro-simulons !

# Les grands panels de la statistique publique

1. Enquête emploi  
+ données fiscales
  2. EDP rénové  
+ panel DAS/DADS + ...
  3. EPAS ou EPIBAM  
+ ESPS
- + des panels plus petits en France

# Les Bases de données médico-économiques

- Enquête santé INSEE-CREDOC 1970
  - Appariement des données d'enquêtes et de sécurité sociale
  - Andrée et Arié Mizrahi « Premiers sondages français dans les dossiers de sécurité sociale et appariement avec les enquêtes auprès des ménages », in Lavallée et Rivest (dir) « Méthodes d'enquêtes et sondages. Pratiques européenne et nord-américaine » Actes du 4ème colloque Sondages de la SFdS, Dunod, Paris, 2006
- Puis ... 2002, 2009, avec le Nir (n° de sécu)

# Les Bases de données médico-économiques

- Depuis 1976, l'EPAS
  - apparié depuis 1988 avec l'ESPS
- Depuis 2004 le SNIIR-AM
  - apparié avec le PMSI (avec FOIN)
- Maintenant l'EPIB-AM
  - bientôt apparié avec ESPS,
  - et avec données de mortalité,
  - et avec carrières CNAV (avec Plastico)

# Les Bases de données médico-économiques

- François Lenormand
  - «Le système d'information de l'assurance maladie», Le Courrier des statistiques, n° 113-114, 2005

# Les Bases de données médico-économiques

- Prochainement Constances
  - 200 000 personnes
  - données appariées issues de :
    - centres d'examen de santé cetaf
    - SNIIR-AM et PMSI
    - carrières CNAV
    - RNIPP-Cepidc (mortalité)

# Les Bases de données médico-économiques

- Goldberg M., Quantin C., Guéguen A., Zins M.
  - “Bases de données médico-administratives et épidémiologie : intérêts et limites », Le Courrier des statistiques, n° 124, 2008.

# Enquête et fichiers de l'emploi

- Deux familles de données bien complémentaires
  - Enquête Emploi
    - concept BIT, mais pas source locale
  - Sources administratives DADS+UNEDIC
    - information locale, mais pas concept BIT
- Chroniques 8 du CNIS «emploi, chômage, précarité»

# Enquête et fichiers de l'emploi

- Deux familles de données bien complémentaires
  - Enquête Emploi
    - concept BIT, mais pas source locale
  - Sources administratives DADS+UNEDIC
    - information locale, mais pas concept BIT
- Chroniques 8 du CNIS «emploi, chômage, précarité»
- Marions-les!

# Appariement

## Enquête et fichiers de l'emploi

---

- Quels bénéfices ?
- Quelles difficultés ?

# Appariement enquête/fichiers de l'emploi

## Quels bénéfices ?

- Pour l'enquête emploi : données DADS surtout
- Pour les fichiers administratifs
  - Correction BIT à partir de données individuelles (niveau national)
  - Transférer le concept BIT au niveau local (zone d'emploi)
- Pascal Ardilly « Panorama des principales méthodes d'estimation sur les petits domaines », INSEE, Méthodologie statistique, n° M0602, 2006, cours du CEPE

# Appariement enquête/fichiers de l'emploi

## Comment ?

- Sur le 1/6ème sortant
- Questions subsidiaires :
  - Appariement rétrospectif (5 ans) ?
  - Poursuivre l'appariement après les 6 trimestres ?
- Quelles difficultés ?
  - Comme pour l'enquête décennale santé :
    - Nir + structure ménage
  - Mais, Comme pour l'enquête décennale santé :
    - Thème très prioritaire et encouragements de la CNIL
- Méthode ?

## Méthode : Engager une réflexion collective

- 1999 groupe de travail SFdS-INSEE-CNIS-CNIL
  - « Transcription en droit français de la directive européenne du 24 octobre 1995 »  
Rapport du CNIS n° 55, 2000.
- 2009 : Concertation à reprendre pour les appariements

# Méthode

- Catherine Quantin propose :
  - hachage du Nir avec clé unique pour l'épidémiologie,
  - puis cryptage avec clé spécifique à l'étude
- Pourquoi pas :
  - Hachage du Nir avec clé unique pour la stat publique
  - et cryptage spécifique ... avant transmission à l'INSEE
  - plutôt que suppression du Nir

# Les données longitudinales dont les économistes rêvent

- Biographies professionnelles et salariales des membres du ménage
- Biographies familiales tout au long du cycle de vie
- Biographies résidentielles
- Biographies patrimoniales
- (Panel des consommations individualisées)

Et le tout pour les mêmes ménages !

# Choix joints au sein du ménage

## Trois exemples d'axes de recherche

- Interactions entre travail, résidence, investissement
- Investissements et transferts intra-familiaux/sociaux
- Interactions entre mariage, fécondité, travail

# Premier axe de recherche : interactions travail, résidence, investissement

- localisation du logement,
- Achat/location,
- Investissements immobilier/financiers,
- Localisation de l'emploi de chaque conjoint,
- Salaire et autres ressources, horaires pour chaque conjoint
- Autres activités
  - domestiques/loisir
  - en commun/assignables/substituables

# Second axe de recherche : interactions Investissements/transferts intra-familiaux & sociaux

- éducation des enfants,
- héritage, donations,
- pensions alimentaires,
- assurance-vie,
- cotisations retraite,
- autres investissements immobilier/financiers,
- Impact et rôle de la législation

## Troisième axe de recherche : interactions mariage / fécondité / travail

- Trajectoires après divorce,
- Concilier travail/fécondité,
- retraites et pension de réversion,
  - Anticipation ?
  - Prise en compte du conjoint ?
- Législation & politiques économiques préconisées ?

# Ce rassemblement de données existe-t-il ?

- Oui aux Etats-Unis :
  - PSID : Panel ménage, cinquantenaire (3 générations)
  - NLSY : Suivi d'une génération de jeunes
- Oui au Royaume Uni (BHPS) et en Allemagne (GSOEP)
- En France
  - avortement du panel lorrain sur ce modèle
  - Interruption du panel européen
  - Autres sources ?

## En France, autres sources ?

- Panels ménage de court terme (enquête emploi, SILC)
- Appariements de fichiers administratifs
- (DADS, ASSEDIC-ANPE, CNAV)
- Echantillon Démographique Permanent
  
- Mais quelle unité : ménages ou individus ?

# En France, autres sources ?

- Rappel de l'objectif :
  - suivre à long terme chaque individu dans ses décisions jointes au sein du ménage
- Les appariements envisageables peuvent-ils être étendus au ménage ?
- Si NON, associer panels et appariements
- Micro-simulations ?
  - Apport limité pour étudier les interactions
  - Intérêt pour simuler différentes politiques
  - Compléments plutôt que substituts à l'appariement