

# Données ouvertes en santé

Pierre-Yves Geoffard

*Ecole d'Économie de Paris*



# Statistique publique

- Tendances de fond : utilisation des données de gestion à des fins statistiques
  - Recommandations du CNIS
  - Complément des données d'enquête
    - Alléger la collecte
    - Eviter les redondances
  - Mais difficultés
    - Systèmes de gestion -> extraction -> transformation + documentation -> base de données
    - Protection de la vie privée / du secret des affaires
    - Cadre légal et réglementaire touffus (rapport Bozio-Geoffard)



# Potentiel des « données de gestion »

- Grande taille (exhaustivité)
- Coût de préparation < coût de collecte
- Mais:
  - Bcp d'observations, peu de variables (+ si appariement)
  - Qui doit payer le coût de préparation ?
  - Dépend de l'utilité pour la gestion publique (ex. K tax Suède)
  - Accès (recherche) régi par des dispositifs différents:
    - Données fiscales
    - Données de santé
    - Autres données
  - Statistique publique + « données de caisses » (Sécurité sociale)



# Potentiel des « données de gestion »

- Enjeux éthiques, protection de la confidentialité
- Arbitrage « big » contre « open »
  - Données agrégées / données individuelles
  - Anonymisation / identifiant (suivi dans le temps, appariement)
  - Risque de réidentification
  - Accès sécurisé : à distance (CASD) ou pièce fermée (Banque de France)



# Potentiel pour l'évaluation

- Description
  - Notamment focalisation sur des sous-populations
    - Parfois difficile à enquêter (non réponses...)
    - Très utile lorsque forte concentration / hétérogénéité
      - Ex: top 1%
  - Evolution temporelle, spatiale
    - Si stabilité du dispositif
- Evaluation d'impact (ex post)
  - Besoin de variations (traitement / contrôle)
    - « le plus exogène possible »
    - RCT, « expériences naturelles »
    - Tirage aléatoire, seuils (RDD)
- Evaluation ex ante (micro-simulations)



# Données de santé

- SNDS

- SNIIRAM + PMSI
- Et bien plus! (données de gestion hospitalière,...)
- Uniquement sur le RO
  - Dépense de soins
  - Remboursement par la SS
  - Rien sur les complémentaires
  - Peu “médicalisées” (pas de codage des diagnostics en ville)



# Quelques exemples / données Françaises

## 1. Description

- Medical Spending in France: Concentration, Persistence and Evolution before Death, avec C. Gastaldi-Ménager, G. de Lagasnerie, *Fiscal Studies*, 2016, n° 37 (3-4), pp 499-526.
  - SNIIRAM (plus précisément EGB)
- The subjective value of a life with Down syndrome: Evidence from amniocentesis decision", avec T. Gadjos & C. Garrouste, *Journal of Economic Behavior and Organization*, 2016, (127), pp 59-69.
  - base hospitalière portant sur les décisions d'IMG

## 2. Evaluation

- Réformer le système de remboursement des soins de ville : une analyse par microsimulation" avec G. de Lagasnerie, *Economie et Statistique*, 2012, n°455-456, pp 89-112
  - microsimulations, toujours l'EGB mais apparié avec l'enquête santé
- Workplace smoking ban effects on unhappy smokers" avec Cl. de Chaisemartin & A.-L. Le Faou, *Health Economics*, 2011, 20(9), pp. 1043-1055.
  - expérience quasi-naturelle : évaluation de l'impact de la loi / tabagisme au travail, à partir de données recueillies par des services hospitaliers d'aide au sevrage tabagique

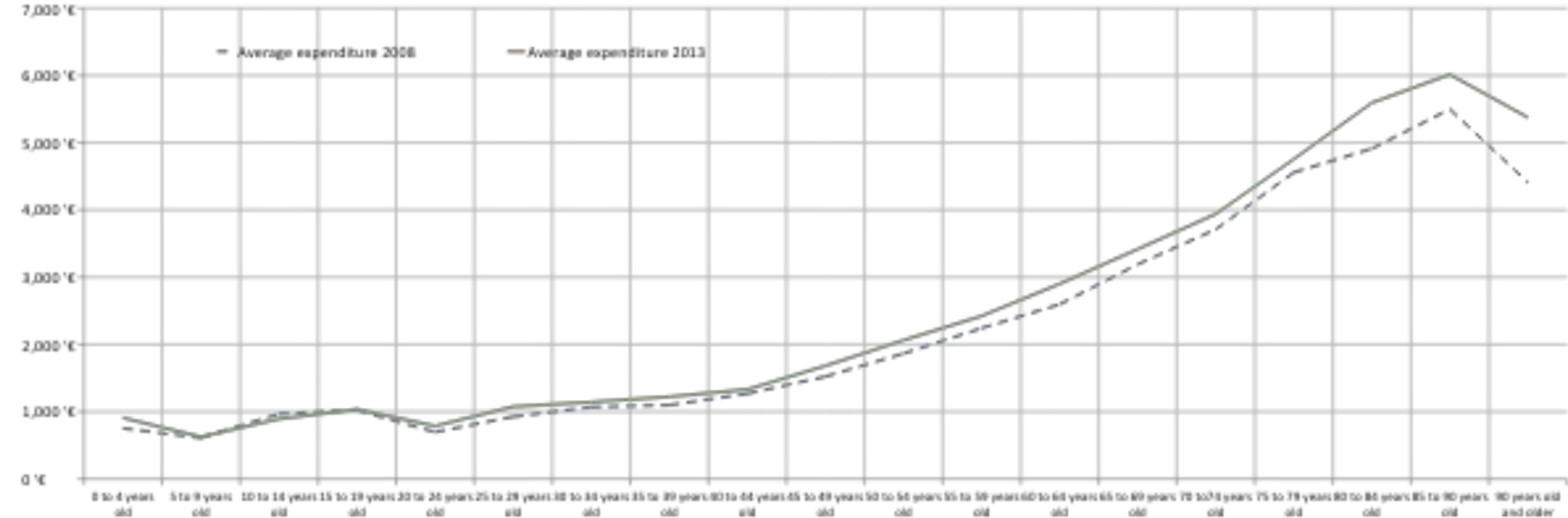


# 1. Dépense de soins

- Projet de comparaison internationale des dépenses de soins
  - Distribution de la dépense (micro), beyond OECD
  - Concentration de la dépense, persistance, dépenses avant décès
  - Fiscal Studies (2016)
- Travail joint avec deux collègues de la CNAMTS
  - Accès facilité + bonne connaissance des bases
  - Echantillon généraliste des bénéficiaires (Régime Général), 2008-2013
    - Au 1/97ème
    - Anonymisé
    - 500 000 individus



# Average expenditure by age (2008, 2013)



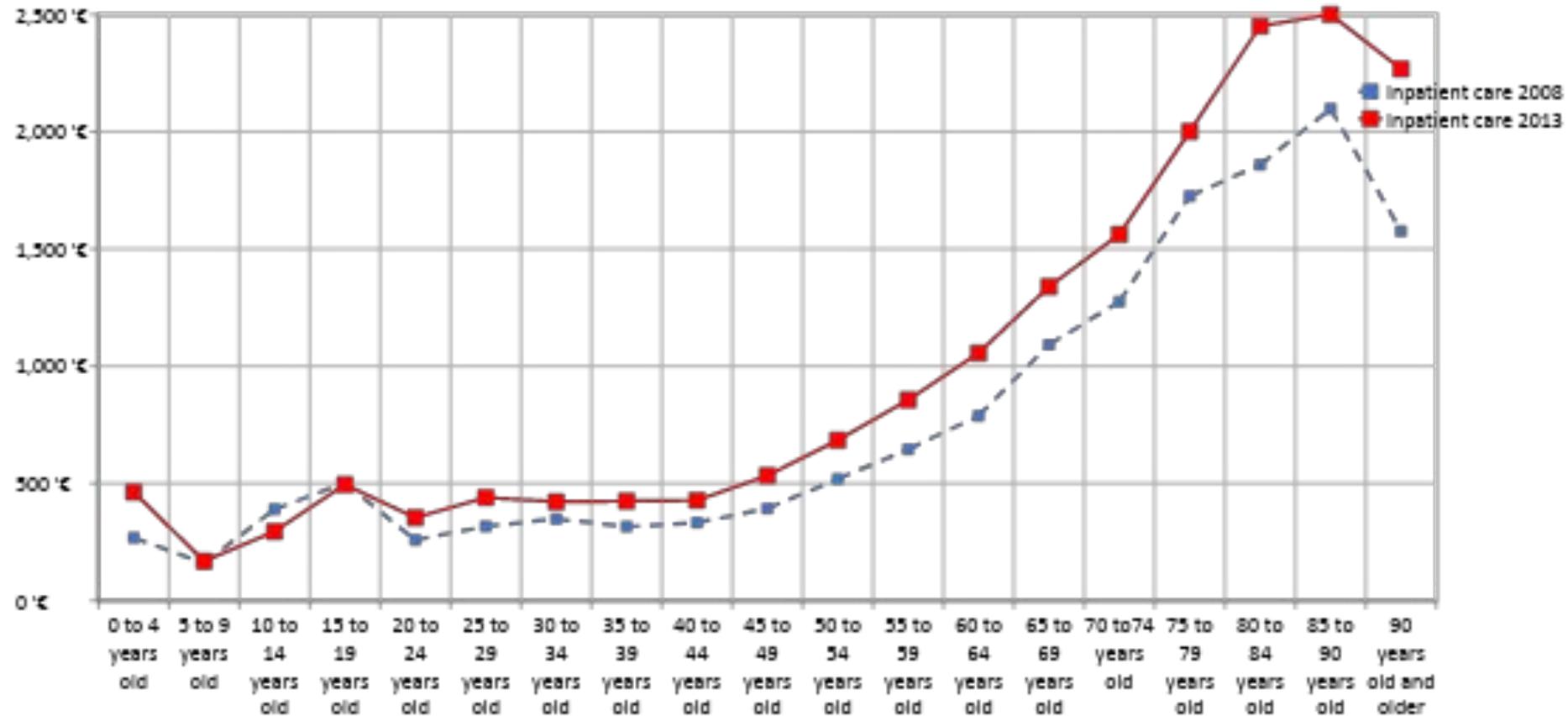
Source : SNIIRAM/EGB-PMSI, 2008-2009-2010-2011-2012-2013

March 2015



# Average care expenditure by age and type of care (2008, 2013)

## Inpatient care



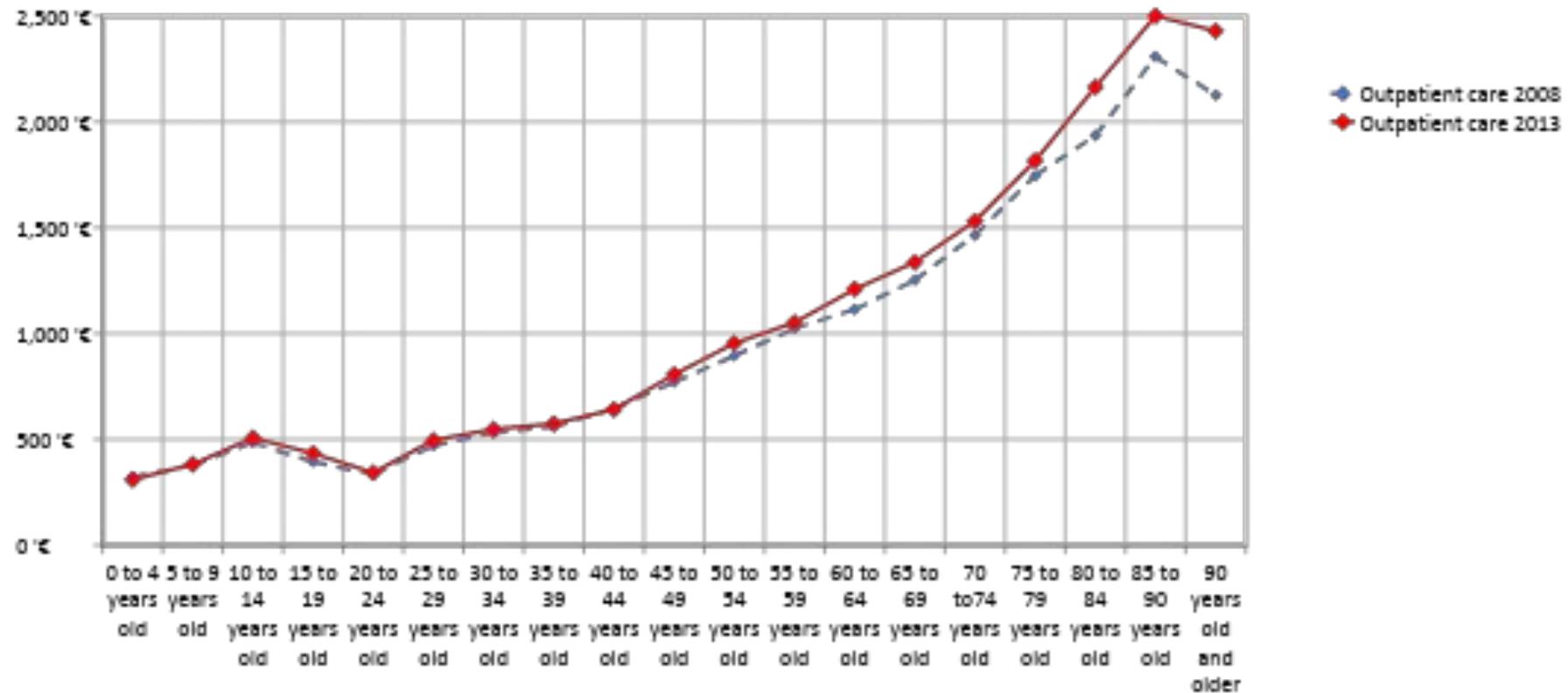
Source : SNIIRAM/EGB-PMSI, 2008-2009-2010-2011-2012-2013

March 2015



# Average care expenditure by age and type of care (2008, 2013)

## Outpatient care



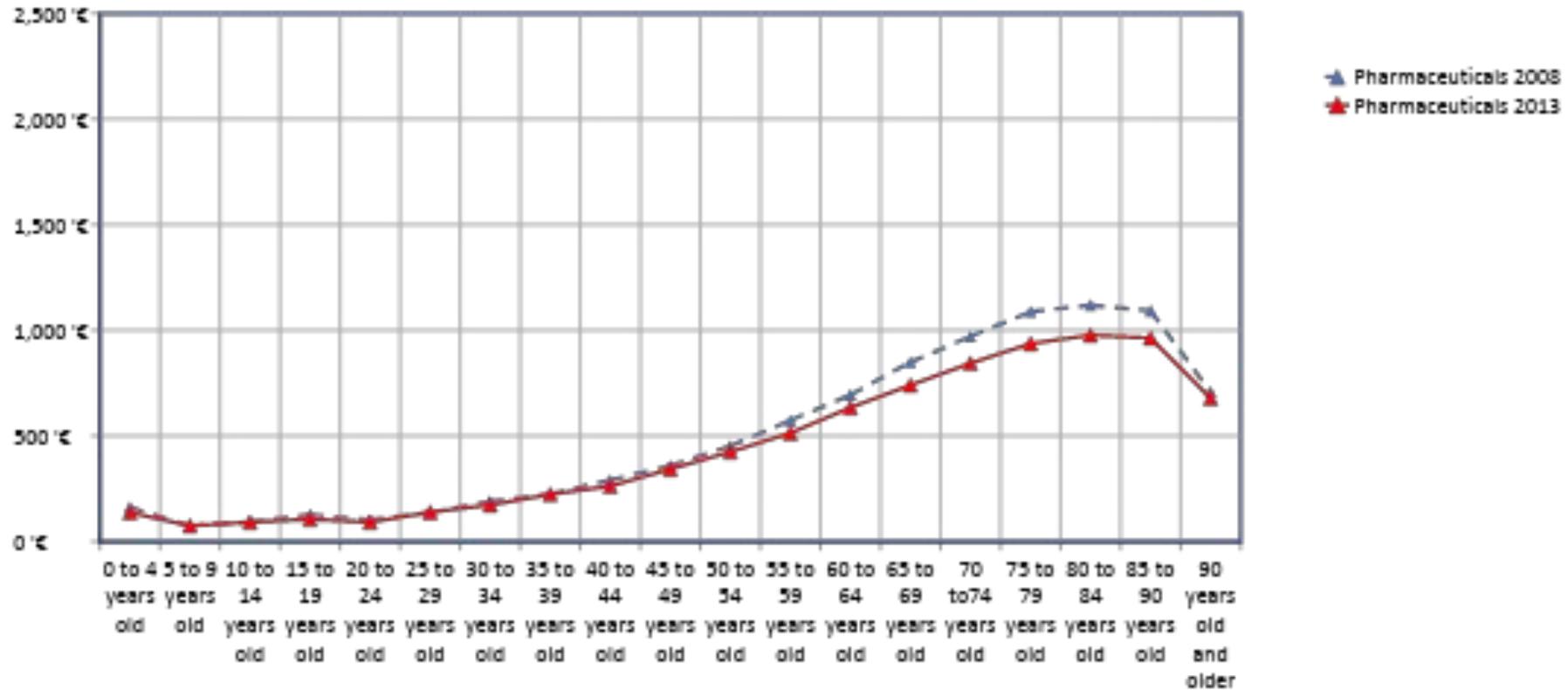
Source : SNIIRAM/EGB-PMSI, 2008-2009-2010-2011-2012-2013

March 2015



# Average care expenditure by age and type of care (2008, 2013)

## Drugs



Source : SNIIRAM/EGB-PMSI, 2008-2009-2010-2011-2012-2013

March 2015



# Concentration des dépenses

- Chaque année, 10% de la population concentre 62% de la dépense de soins
  - Diminue avec l'âge
  - Augmente sur la période
  - Plus forte concentration pour les remboursements (par le RO) que pour la dépense
    - Indique que l'assurance joue un rôle...
- Moins forte concentration pour la dépense / 5 ans



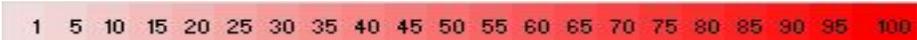
# Persistance des dépenses

- Matrice de transition de décile à décile
  - Ensemble de la population
  - Population en ALD
- Grande taille de l'échantillon: analyse possible aussi par centile



# Total medical expenditure transition matrices in the all population: decile

Scale:



Decile in 2008

Decile in 2009

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	56	14	8	6	5	4	3	2	2	1
D2	19	27	18	12	8	6	4	3	2	1
D3	9	20	21	16	11	8	6	3	3	2
D4	6	13	18	18	15	11	8	5	4	2
D5	3	8	13	17	18	14	11	8	5	3
D6	3	5	8	12	16	18	15	11	7	5
D7	2	4	6	8	12	16	19	16	10	7
D8	1	2	3	5	8	12	17	23	18	11
D9	1	2	3	4	5	7	11	19	30	19
D10	1	1	2	3	3	5	6	10	21	48

Decile in 2010

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	50	13	8	6	5	4	4	4	4	3
D2	17	26	17	12	9	6	5	3	3	2
D3	10	20	19	15	11	9	6	4	3	2
D4	6	14	17	17	14	10	8	6	4	3
D5	4	10	13	16	17	14	10	7	5	4
D6	3	7	10	13	15	16	14	10	7	5
D7	2	5	7	9	12	16	17	15	10	7
D8	2	3	4	6	8	12	17	21	16	11
D9	1	3	3	4	6	7	11	19	27	19
D10	1	2	2	3	4	5	7	11	21	43

Decile in 2013

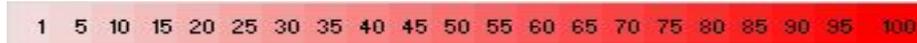
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	45	13	8	6	6	5	4	4	4	5
D2	16	23	16	12	9	7	6	4	4	3
D3	10	19	17	14	11	9	7	6	4	3
D4	7	14	16	15	13	11	9	7	6	4
D5	5	10	14	15	14	12	10	8	7	5
D6	4	8	11	13	14	14	12	10	8	6
D7	3	6	8	10	12	14	15	13	10	8
D8	3	4	5	7	9	13	16	17	15	12
D9	2	3	4	5	6	9	12	18	22	18
D10	2	2	3	4	5	6	8	13	20	37

Source : SNIIRAM/EGB-PMSI, 2008-2009-2010-2011-2012-2013



# Total medical expenditure transition matrices : decile for people with chronic diseases

Scale:



Decile in 2009

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	25	7	7	9	9	10	10	7	7	9
D2	18	19	14	10	8	7	7	5	4	8
D3	10	13	15	14	11	10	8	7	5	8
D4	5	7	13	16	13	12	10	8	7	9
D5	3	4	6	12	18	15	13	10	7	10
D6	2	2	4	6	12	21	18	14	10	11
D7	1	1	2	3	7	14	23	21	14	14
D8	0	1	1	2	3	7	16	29	24	16
D9	0	0	1	1	1	3	7	18	39	29
D10	0	0	0	1	1	2	3	8	21	64

Decile in 2008

Decile in 2010

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	17	5	4	4	5	6	8	12	17	22
D2	14	15	10	9	8	8	8	7	6	14
D3	9	11	13	11	9	9	9	9	8	13
D4	5	7	11	13	12	11	10	10	9	13
D5	3	4	6	10	14	15	13	11	9	14
D6	2	3	3	6	11	18	17	14	11	15
D7	1	1	2	3	7	13	21	19	16	16
D8	1	1	1	2	3	7	16	27	22	20
D9	1	0	1	1	2	3	8	20	35	30
D10	1	0	1	1	1	2	4	9	22	58

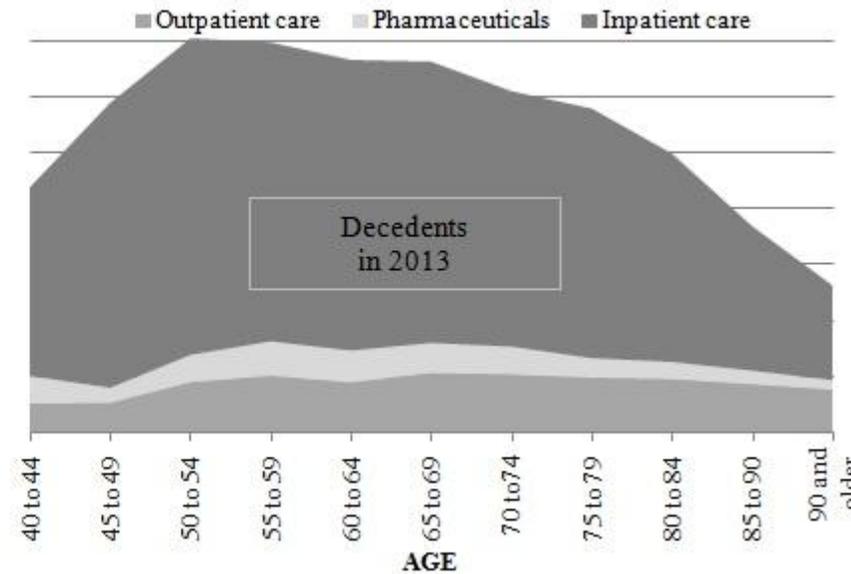
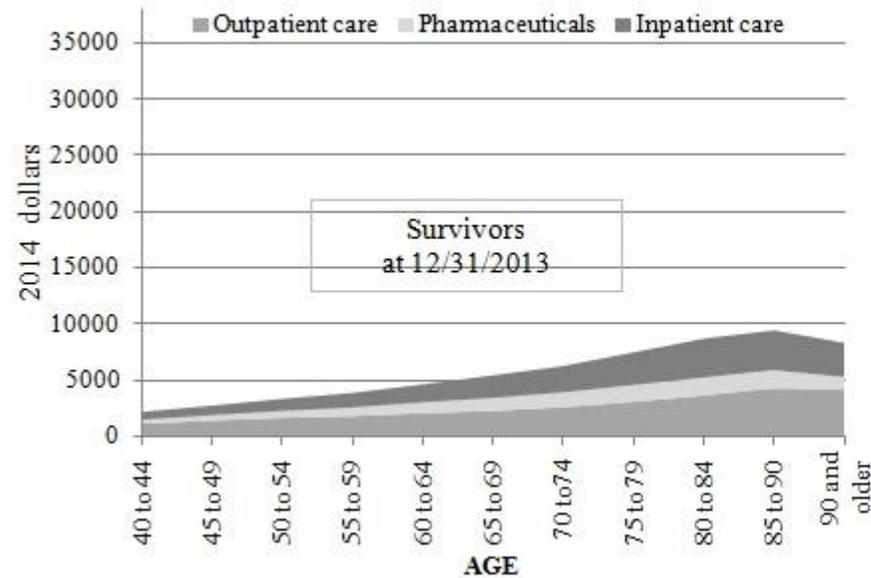
Decile in 2013

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	8	3	3	3	5	7	10	13	17	32
D2	7	8	6	8	8	9	8	9	11	26
D3	5	7	6	7	8	10	10	11	12	24
D4	4	5	6	8	9	10	11	12	13	22
D5	3	3	4	7	10	11	12	12	14	24
D6	2	2	3	6	9	12	13	14	15	23
D7	2	1	2	3	6	12	15	17	17	24
D8	1	1	1	2	4	8	15	20	21	26
D9	1	1	1	1	2	5	10	19	28	33
D10	1	1	1	1	2	3	6	11	22	52



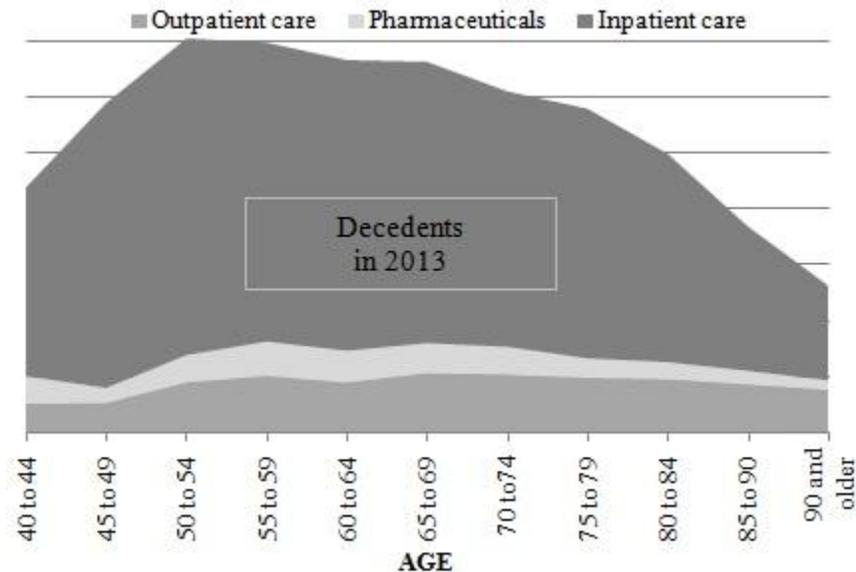
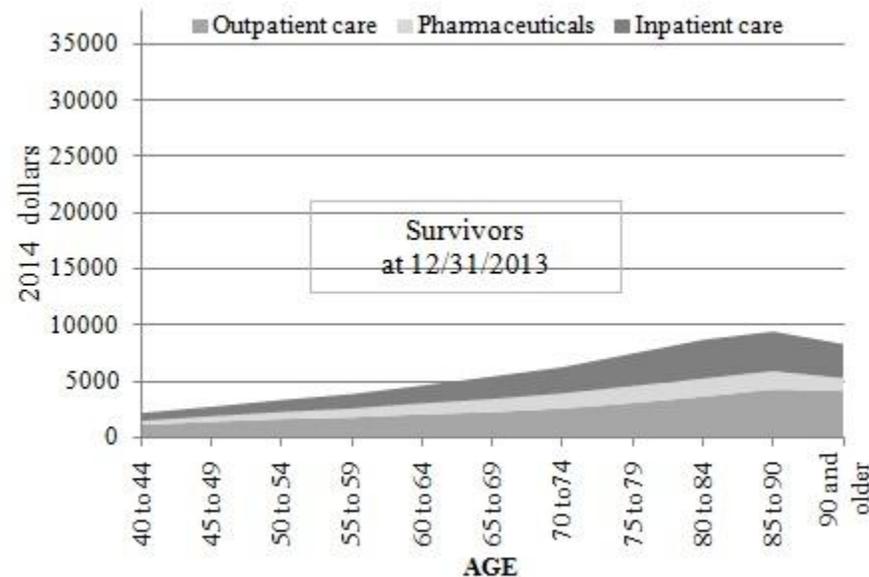
# Dépense avant décès

- Lieu commun : très forte concentration de la dépense de soins sur les derniers mois de la vie
  - “50% de la dépense sur les 6 derniers mois”
- Les données : dépense avant décès en effet bien plus élevée



# Dépense avant décès

- Lieu commun : très forte concentration de la dépense de soins sur les derniers mois de la vie : “50% de la dépense sur les 6 derniers mois”
- Les données : dépense avant décès en effet bien plus élevée
  - Diminue avec l'âge au décès
- Extrapolation : 3 dernières années de vie = 22% de la dépense vie entière



## 2. Décision d'amniocentèse

- Détection d'une mutation génétique (trisomie 21)
- Risque augmente avec l'âge de la mère
  - Evaluation du risque (non invasif)
  - suivi (ou pas) d'une amniocentèse
    - Risque de fausse couche induit par l'amniocentèse
  - Amniocentèse remboursée selon le niveau de risque (seuil 1/250)
  - Arbitrage bénéfice/risque + coût
- Base de données hospitalières
  - Hôpitaux parisiens, 2003-2007
  - 28341 femmes, âge moyen 30 ans, peu de variables socio-éco



## 2. Décision d'amniocentèse

- Méthode : les décisions individuelles révèlent les valeurs relatives (implicites) attribuées à :
  - Une naissance sans trisomie 21
  - Une naissance avec trisomie 21
  - Une absence de naissance

Le seuil de remboursement révèle aussi implicitement la valeur sociale attribuée à ces trois cas



## 2. Décision d'amniocentèse : résultats

Comparaison près du seuil de 1/250 (RDD)

Élasticité prix de la décision d'amniocentèse:

- baisse de 50% de la proba juste au dessus du seuil

La valeur sociale implicite d'un enfant avec trisomie 21 est bien plus faible que celle inférée des décisions individuelles.



# 3. Bouclier sanitaire

- Enquête santé 2003, appariée au SNIIRAM/PMSI
  - 6740 ménages, 15042 individus
  - suffit pour regarder les 10%, pas pour les 1%
- Tickets modérateurs, franchises, dépassements...
- (pseudo)RAC très élevé pour certains
  - avant remboursement par complémentaire
  - l'exo TM / ALD protège mal les plus malades
    - polypathologies

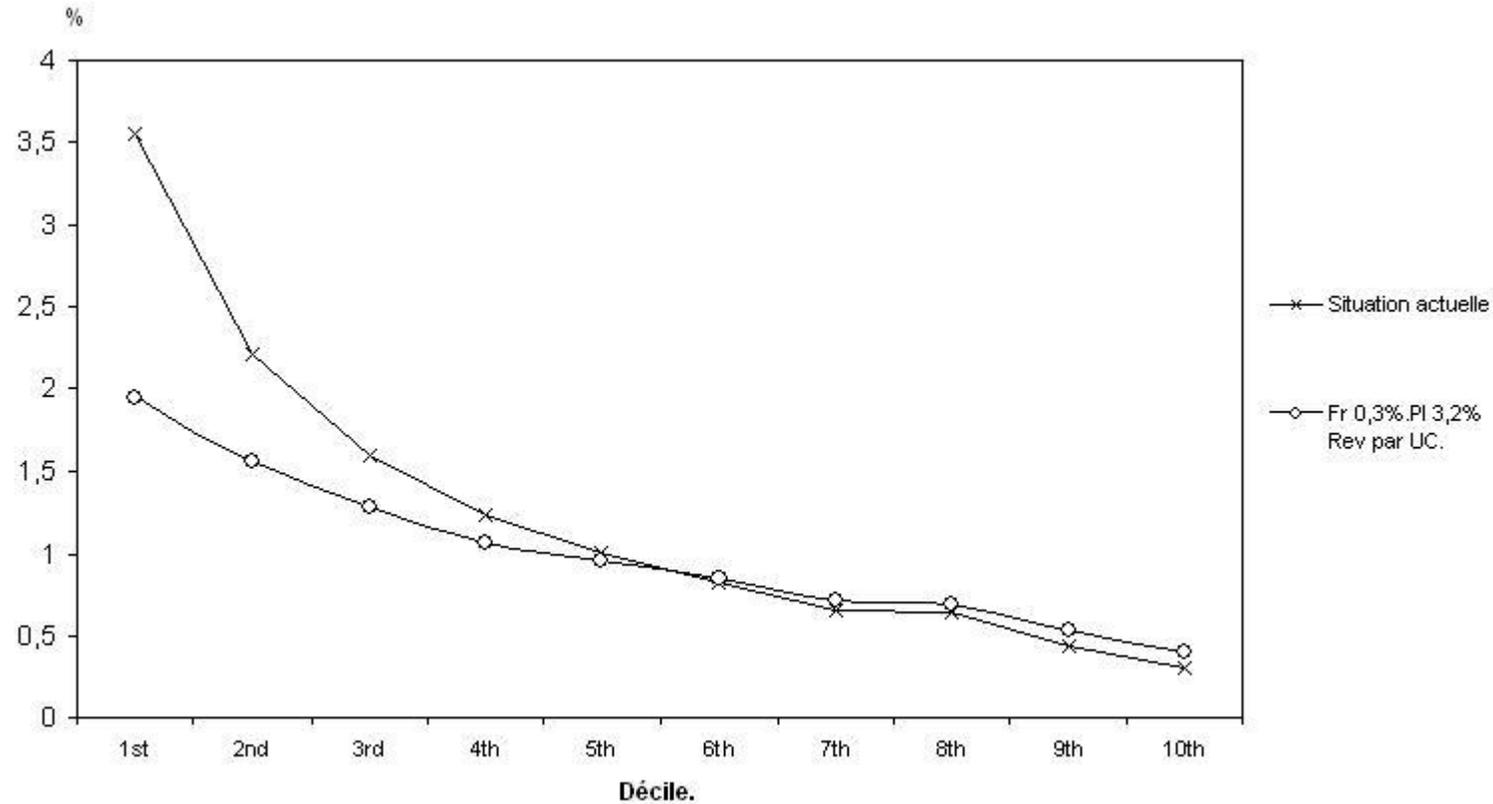


# 3. Bouclier sanitaire : microsimulation

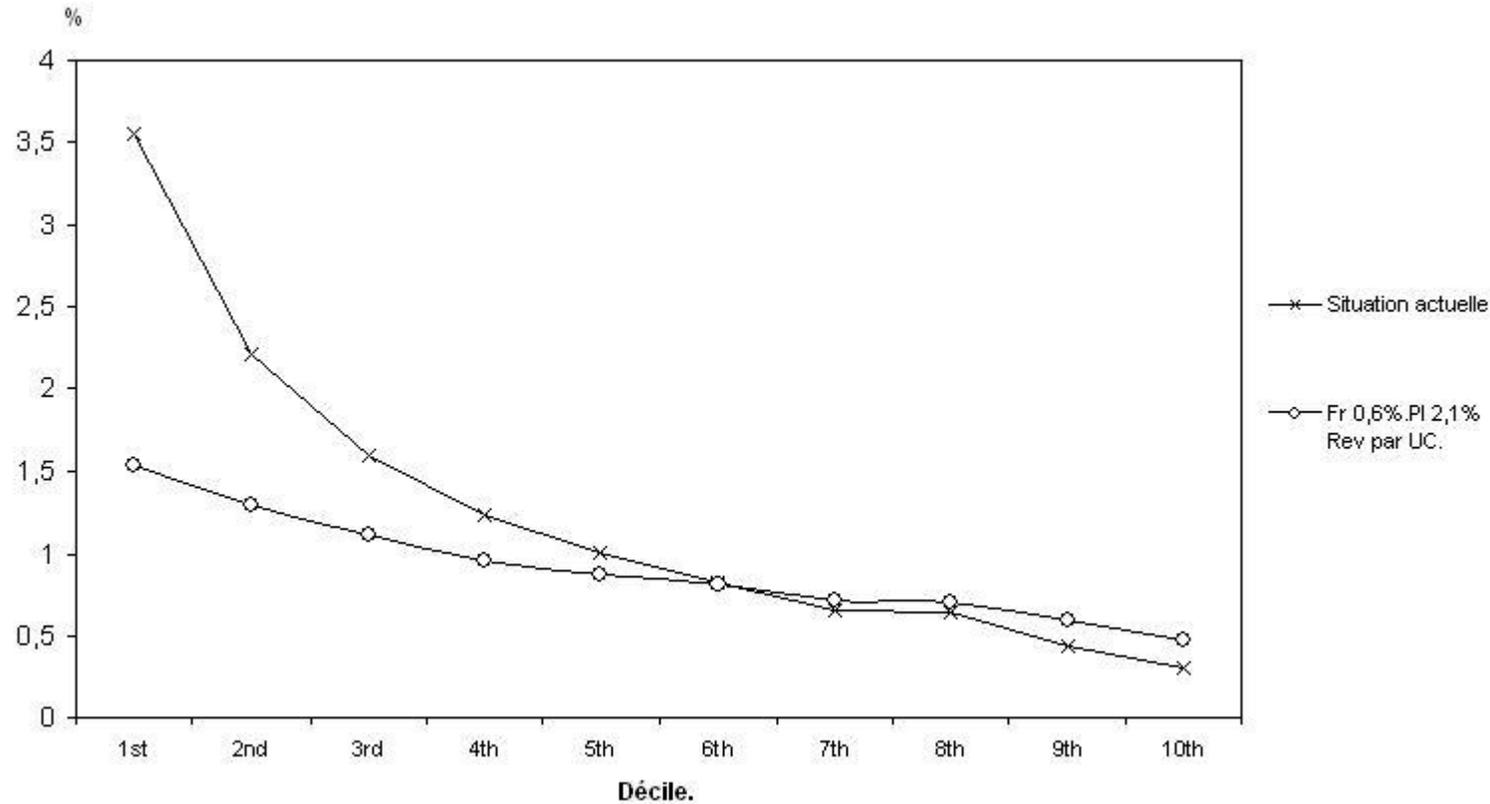
- Simulation d'une réforme du TM:
  - Introduction d'un plafond annuel sur le p-RAC
  - Financement par une franchise annuelle
    - Plafonds et franchises uniformes, ou fonction du revenu (!!)
- Résultats des simulations pour un dispositif fonction du revenu par UC du ménage:
  - Franchise à 0,3% = plafond à 3,2%
  - Franchise à 0,6% = plafond à 2,1%
  - Franchise à 1,2% = plafond à 1,53%
- Part du revenu annuel consacré au TM



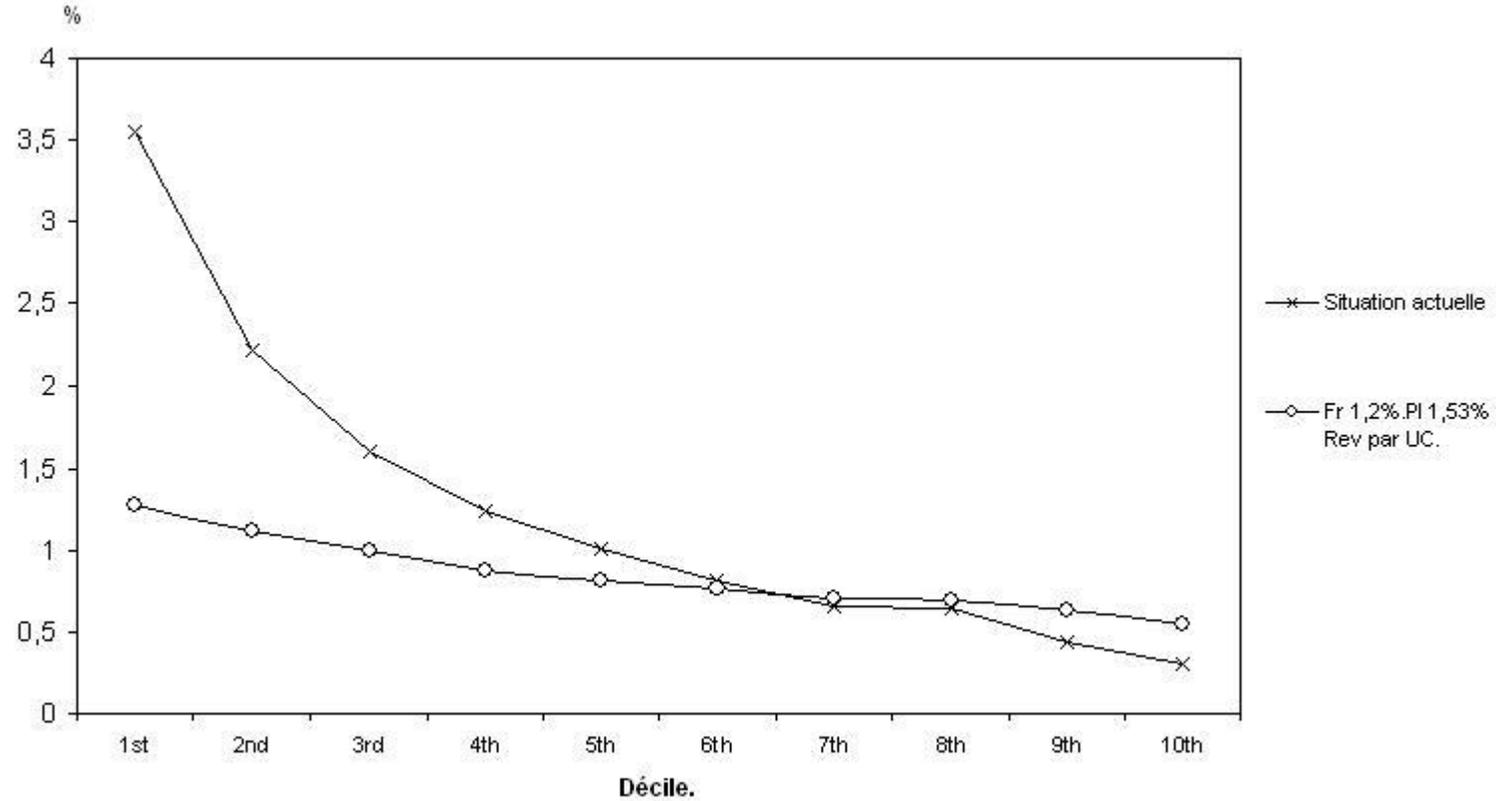
### 3. Bouclier sanitaire : franchise à 0,3%



### 3. Bouclier sanitaire : franchise à 0,5%



### 3. Bouclier sanitaire : franchise à 1,2%



# 4. Loi Tabac de 2006

- Interdiction de fumer au travail
  - Partielle en 1991, peu appliquée
  - Complète en 2006.
- Base de données centralisée à l'HEGP
  - Consultations dépendance tabagique : aide à l'arrêt
  - 27000 patients, 2004 – 2008.
- La loi a-t-elle entraîné une baisse de la consommation?
  - (une augmentation des demandes d'aide)
  - Comparaison avant/après, pour différents groupes (en emploi ou non)
    - Augmentation, surtout dans les zones pluvieuses et froides.



# 4. Loi Tabac de 2006

WORKPLACE SMOKING BAN EFFECTS ON UNHAPPY SMOKERS

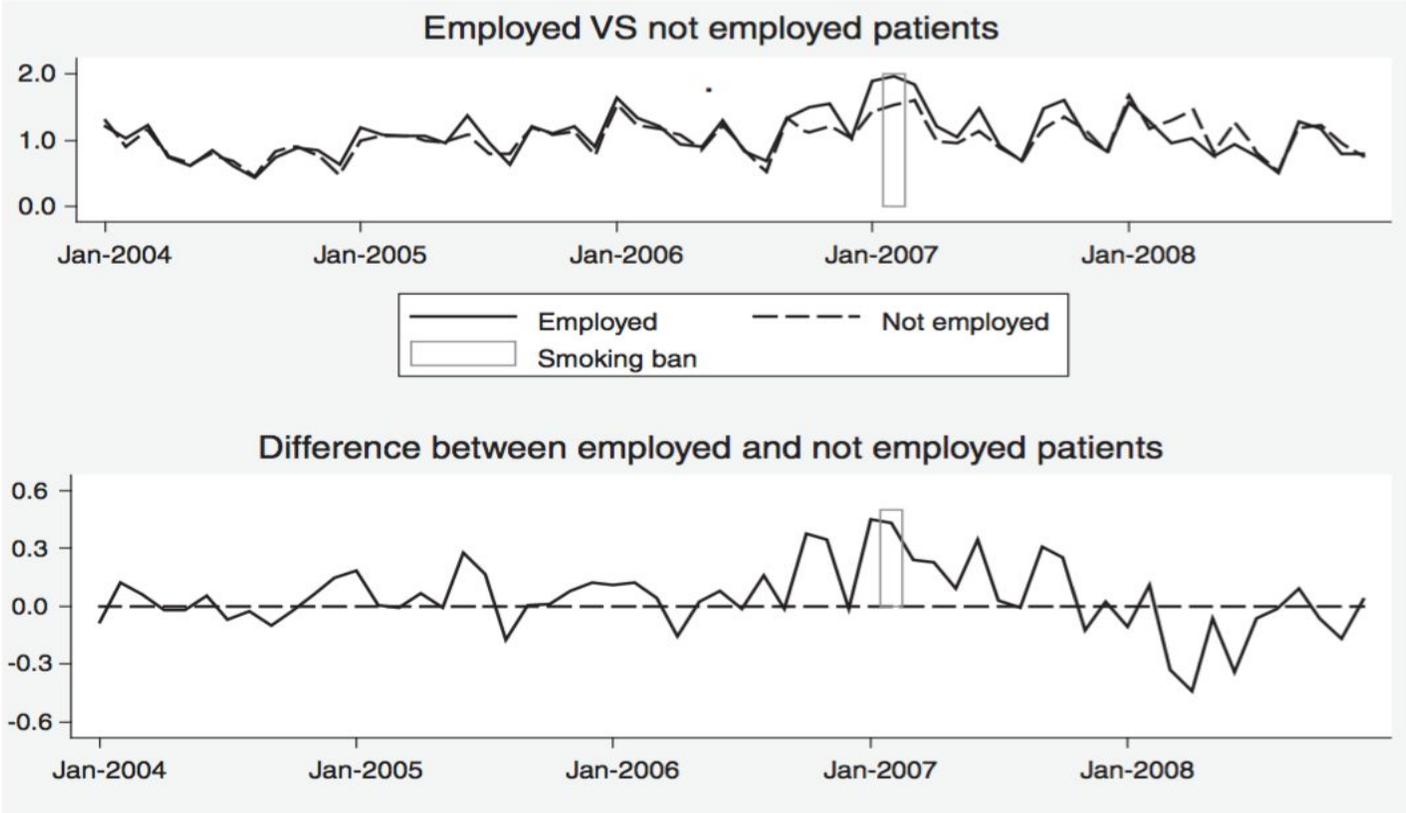


Figure 1. Number of patients visiting cessation services



# Conclusion

- L'ouverture des données de santé permet de conduire des évaluations
  - Imparfaites :
    - peu de variables socio-éco
    - Pas toujours la bonne variable d'outcome
  - Mais moins coûteuses, et plus rapides qu'un RCT...
- Les données du SNIIRAM/PMSI
  - Exceptionnelles par la taille (65 millions d'individus...)
  - Mais pas par le contenu (peu d'infos cliniques, socio-éco,...)
- Au delà du SNIIRAM
  - Données hospitalières de gestion (entrepôt de données de l'APHP,...)
  - Suivi de cohortes (finalité clinique)
- Accès, coût de préparation / documentation (English please)
- Appariement (SNDS, données fiscales, EIC, EIR,...)

