

Calcul d'indicateurs précoces

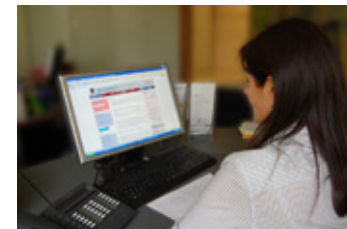
Que peut-on faire ? Et dans quel but ?

Cédric Houdré
Chef de la Division revenus et
patrimoine des ménages

Juliette Stehlé
Division Etudes sociales



Mesurer pour comprendre



04/11/2014

Un problème difficile: **Précocité** vs **Qualité**

Délais de mise à disposition des données monétaires

- Décembre N+1 pour les sources fiscales
- Avril N+2 pour les sources sociales
- Publication en **septembre N+2** du taux de pauvreté monétaire, **+20 mois** après la période de référence

Améliorations récentes sur le non monétaire

- Finalisation des données non monétaires de FR-SILC2013 en mars 2014, +9 mois après la collecte
- Publication par Eurostat de l'indicateur « Severe material deprivation in 2013 » en juin 2014

Si les revenus étaient mesurés en face à face avec SILC, le taux de pauvreté pourrait être publié 9 mois après la collecte

Un problème difficile: Précocité vs **Qualité**

Un choix déjà ancien de recourir aux sources administratives

- Depuis toujours pour les revenus fiscaux (ERF1970-1990, ERF1996-2004)
- Depuis 2005 pour mesurer les revenus sociaux (ERFS2005-2012)
- Depuis 2007 pour l'enquête FR-SILC européenne

Qualité indirecte

Un allègement de la charge d'enquêtes pour les ménages

- Baisse du temps de collecte
- Concentration des questionnaires sur des questions moins sensibles

Qualité directe

Des revenus mieux mesurés

- Revenus mixtes (petites retraites et salaires etc.)
- Revenus peu fréquents (capitaux mobiliers, fonciers)
- Compléter par imputation pour revenus non soumis à l'impôt

Un problème difficile: Précocité vs **Qualité**

Des différences sur la mesure de la pauvreté...

Expérience sur SRCV2005 (L. Dauphin (2008))

Tableau 6 : Taux de pauvreté en fonction de la source utilisée

	13 447 individus ²¹			
	Enquête avant imputation	Enquête après imputation	DGFIP*	DGFIP**
Taux de pauvreté	18,1 %	13,0 %	12,5 %	13,4 %
Seuil de pauvreté ²²	9 037	9 434	9 513	9 551

**Salaires, retraites et allocations chômage sont ceux déclarés auprès de l'administration fiscale.*

Les impôts, les bénéfices des indépendants, les revenus fonciers, les minima sociaux et les allocations sont ceux de l'enquête.

***En intégrant les impôts récupérés auprès de l'administration fiscale.*

Un problème difficile: Précocité vs **Qualité**

... et sur les positions dans la distribution

Pourcentage (pondéré)		Données d'enquête										
		1 ^{er} décile	2 ^e décile	3 ^e décile	4 ^e décile	5 ^e décile	6 ^e décile	7 ^e décile	8 ^e décile	9 ^e décile	10 ^e décile	Total
Données fiscales	1 ^{er} décile	6,4	1,2	0,5	0,4	0,2	0,4	0,4	0,1	0,3	0,1	10,0
	2 ^e décile	1,1	6,1	0,8	0,6	0,7	0,2	0,2	0,2	0,2	0,1	10,0
	3 ^e décile	0,8	1,6	5,6	0,9	0,3	0,3	0,2	0,2	0,1	0,0	10,0
	4 ^e décile	0,7	0,4	2,0	5,1	1,0	0,3	0,2	0,0	0,2	0,0	10,0
	5 ^e décile	0,1	0,2	0,4	1,9	5,4	1,0	0,4	0,3	0,3	0,2	10,0
	6 ^e décile	0,1	0,1	0,3	0,6	1,8	5,3	0,9	0,4	0,3	0,2	10,0
	7 ^e décile	0,2	0,1	0,3	0,2	0,2	1,6	5,7	1,2	0,3	0,4	10,0
	8 ^e décile	0,2	0,2	0,1	0,1	0,2	0,7	1,2	6,0	1,0	0,3	10,0
	9 ^e décile	0,2	0,1	0,1	0,2	0,0	0,2	0,5	1,5	6,2	1,1	10,0
	10 ^e décile	0,2	0,1	0,0	0,0	0,2	0,1	0,2	0,3	1,1	7,7	10,0
	Total	10,0	10,0	10,0	10,0	10,0	10,0	10,0	10,0	10,0	10,0	100,0

40% des individus ne sont pas dans le même décile

Timeliness

Que peut-on faire? Dans quel but ?

Peu de marges dans l'accès aux données administratives

- 4ème role fiscal (quasi-définitif, couvre 95% des déclarations totales)
- Délais de construction des données de la Cnaf

Tirer parti de la modélisation

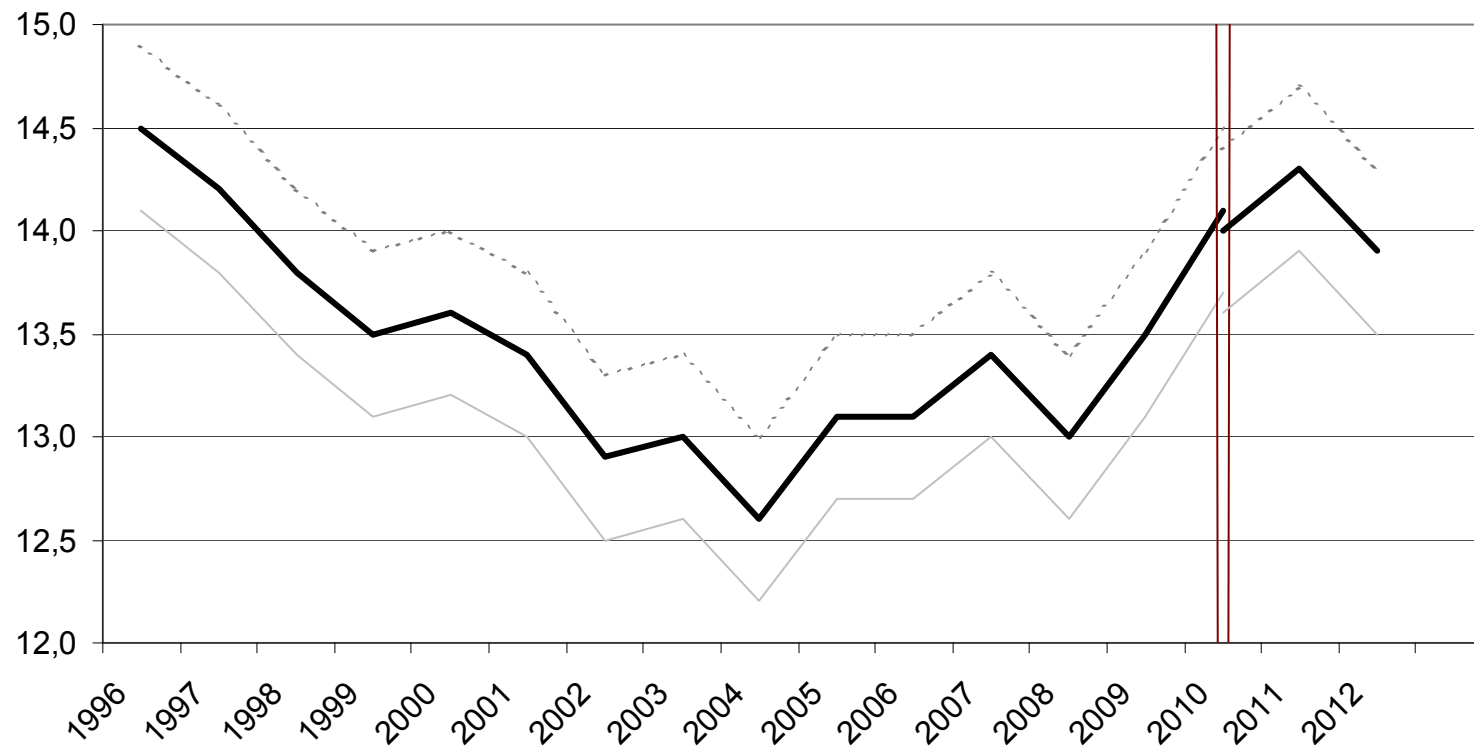
- Approche économétrique
- Approche par microsimulation

Timeliness

Que peut-on faire? Dans quel but ?

Le taux de pauvreté est un indicateur complexe

- les tendances de moyen-terme suivent les inflexions de la conjoncture économique

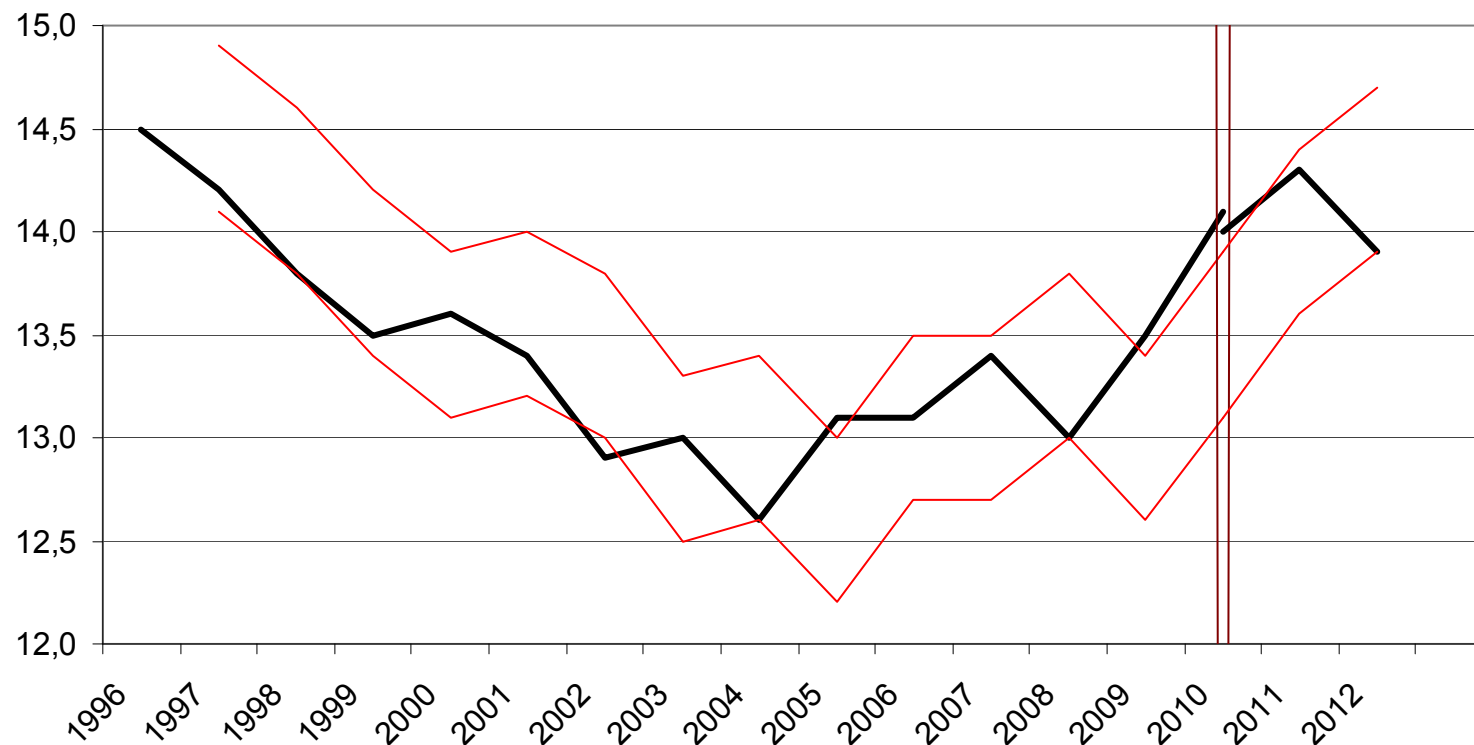


Timeliness

Que peut-on faire? Dans quel but ?

Le taux de pauvreté est un indicateur complexe

- Mais une conjoncture annuelle de la pauvreté reste difficile à établir

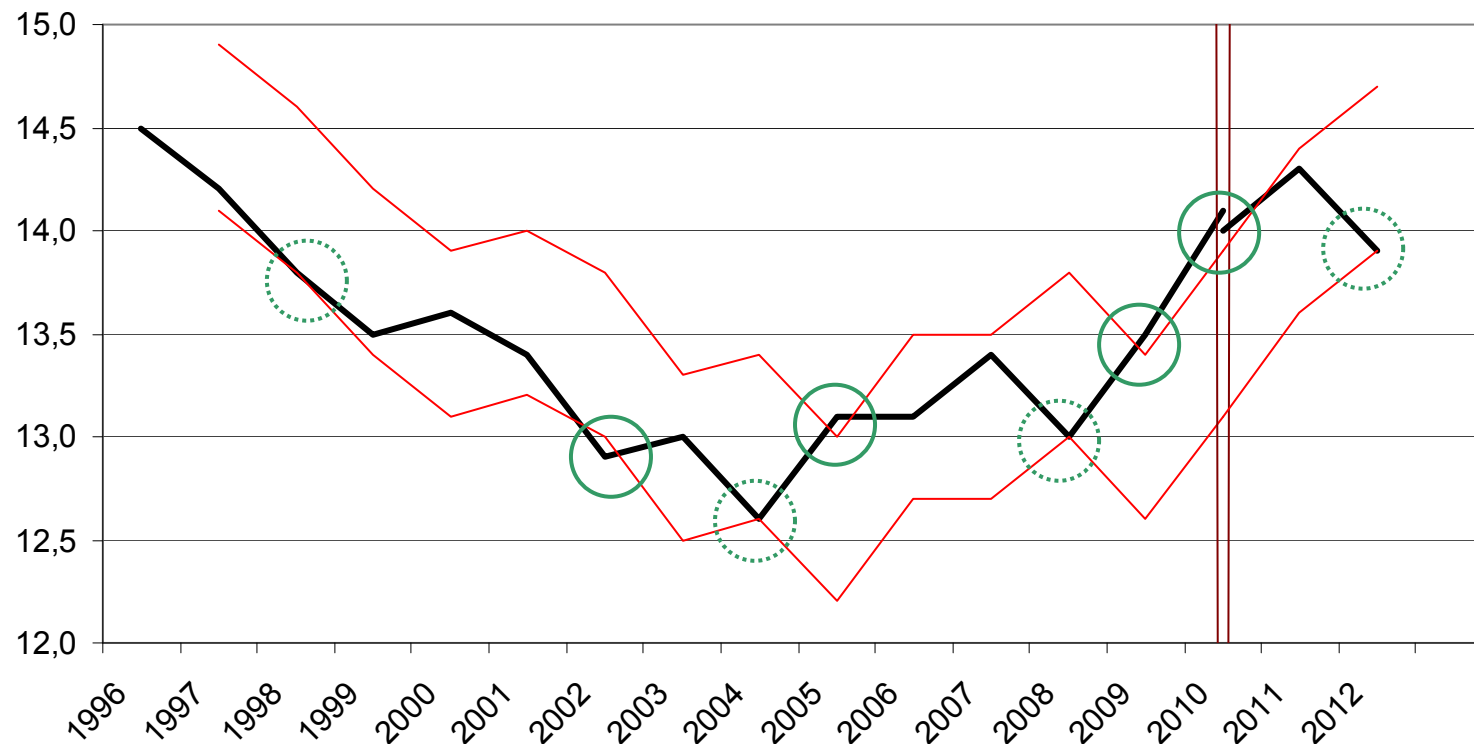


Timeliness

Que peut-on faire? Dans quel but ?

Le taux de pauvreté est un indicateur complexe

- Mais une conjoncture annuelle de la pauvreté reste difficile à établir les évolutions annuelles étant souvent non significatives

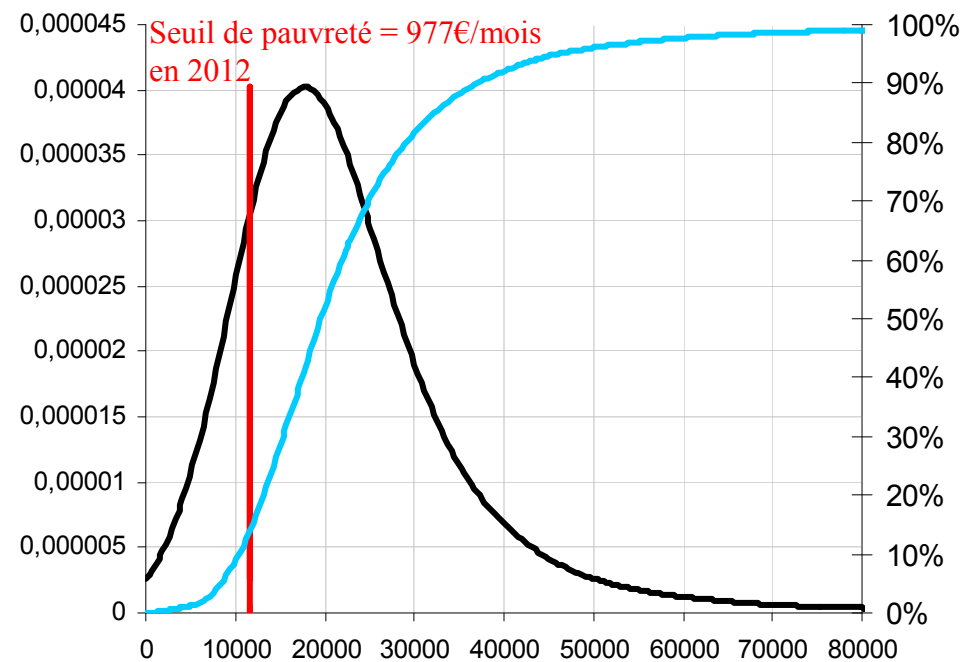


Timeliness

Que peut-on faire? Dans quel but ?

Le taux de pauvreté est un indicateur complexe

- paradoxalement
- sensible à de petites variations du niveau de vie médian et de la concentration autour du seuil



+/-1% du seuil de pauvreté
= +/-10€ / mois
= +/-2,6% du taux de pauvreté
= +/-0,33 point de taux de pauvreté

Timeliness

Que peut-on faire? Dans quel but ?

Le taux de pauvreté est un indicateur complexe

- À la fois « relativement » stable depuis 16 ans
- Et pourtant sensible à de petites variations du niveau de vie médian et de la concentration autour du seuil

But atteignable (?)

Prédire correctement le **sens de variation**

en cas de choc macro majeur

Deux stratégies testées ou en cours...

1. Modélisation économétrique

2. Microsimulation (J. Stehlé)

...

Et une troisième voie qui pourrait être examinée

3. Déclarations pré-remplies

Prédiction **économétrique** d'indicateurs d'inégalités et de pauvreté

Données

Statistiques sur les ressources et les conditions de vie

- 11000 ménages
- panel rotatif sur 9 ans
- appariement sources admin depuis revenus 2007 (collecte 2008)
- collecte en face à face du « revenu mensuel total le mois précédent la collecte »

Prédiction économétrique d'indicateurs d'inégalités et de pauvreté

Principe 1: régression linéaire

$$E [NV_{it}] = RM_{it} \cdot \alpha + NV_{it-1} \cdot \beta + X'_{it} \cdot \gamma$$

Prédiction économétrique d'indicateurs d'inégalités et de pauvreté

Principe 1: régression linéaire

$$E [NV_{it}] = RM_{it} \cdot \alpha + NV_{it-1} \cdot \beta + X'_{it} \cdot \gamma$$

Problème:

distribution prédite trop « tassée » vers la moyenne

Prédiction économétrique d'indicateurs d'inégalités et de pauvreté

Principe 1: régression linéaire

$$E [NV_{it}] = RM_{it} \cdot \alpha + NV_{it-1} \cdot \beta + X'_{it} \cdot \gamma$$

Principe 2: régression quantile non conditionnelle

$$Q_{\tau} [NV_{it}] = RM_{it} \cdot \alpha + NV_{it-1} \cdot \beta + X'_{it} \cdot \gamma$$

Firpo, Fortin & Lemieux (2009) *Econometrica*

Prédiction économétrique d'indicateurs d'inégalités et de pauvreté

Principe 1: régression linéaire

$$E [NV_{it}] = RM_{it} \cdot \alpha + NV_{it-1} \cdot \beta + X'_{it} \cdot \gamma$$

Principe 2: régression quantile non conditionnelle

$$Q_{\tau} [NV_{it}] = RM_{it} \cdot \alpha + NV_{it-1} \cdot \beta + X'_{it} \cdot \gamma$$

$$\text{Gini} [NV_{it}] = RM_{it} \cdot \alpha + NV_{it-1} \cdot \beta + X'_{it} \cdot \gamma$$

$$\text{TauxPauv} [NV_{it}] = RM_{it} \cdot \alpha + NV_{it-1} \cdot \beta + X'_{it} \cdot \gamma$$

$$\text{Q9/Q1} [NV_{it}] = RM_{it} \cdot \alpha + NV_{it-1} \cdot \beta + X'_{it} \cdot \gamma$$

Prédiction économétrique d'indicateurs d'inégalités et de pauvreté

Résultats

Modèle	Taux 2009	Taux 2010	Variation 2009-2010	Sens
EU-SILC	12,9	13,3	0,4	+
Modèle linéaire généralisé	12,8	14,5	1,7	+
Modèle linéaire généralisé (réduit)	13,1	14,7	1,6	+
Logit dichotomique	12,6	13,8	1,2	+
Probit dichotomique	12,6	13,7	1,1	+
Logit multinomial	11,2	12,8	1,6	+

- Surévaluation systématique de l'évolution du taux de pauvreté
- Méthode par régression quantile non conditionnelle pas supérieure à simple probit pour le taux de pauvreté

Microsimulation

Microsimulation : principes généraux

- Avoir un échantillon représentatif et de taille suffisante
- Simuler, au niveau d'une observation (ménage, individu, entreprise), tout ce qui dépend entièrement de la législation
- Simuler une évolution temporelle réaliste en tenant compte des évolutions conjoncturelles et/ou démographiques

- Permet de simuler des scénarios alternatifs
 - ⇒ évaluation *ex-post* de politiques publiques, chiffrages *ex-ante*
- Permet de décrire une situation précocement (pas besoin d'attendre les données administratives définitives)
 - ⇒ nowcasting du taux de pauvreté (EUROMOD, Ines)

Microsimulation de systèmes socio-fiscaux

- Echantillon représentatif de la population d'une année $N-K$
 - Souvent ERFS (France) ou SILC (Europe)
 - Échantillon disponible avec retard : méthodes de repondération ou transitions individuelles entre emploi et chômage pour être représentatif d'une année plus récente N
- Mise à jour des revenus non-simulés entre $N-K$ et N
 - Salaires et revenus des indépendants
 - Allocations chômage
 - Retraites
 - Revenus du patrimoine
- Simuler sur barème, au niveau d'un individu / foyer, les prélèvements et prestations inclus dans le niveau de vie en N
 - Impôt sur le revenu
 - Cotisations/contributions sociales et prélèvements sociaux
 - Prestations familiales
 - Minima sociaux
 - Allocations logement

Difficultés inhérentes à la microsimulation

- Utilisation de données macroéconomiques pour des simulations micro
 - On impute à chaque individu l'évolution moyenne de la sous-catégorie à laquelle il appartient (individus de même catégorie sociale, ou percevant le même type de revenus) : les variations individuelles ne sont pas captées (cessation d'activité, prise d'un emploi...), ni même l'évolution de la dispersion à l'intérieur d'une catégorie
- Des informations manquantes pour simuler fidèlement la législation
 - Informations sur le passé insuffisantes (profondeur temporelle)
 - Niveau de détail insuffisant dans les enquêtes (ressources annuelles et non mensuelles, degré de handicap)
- Modèles généralement non-comportementaux
 - Pas de prise en compte des effets des politiques publiques sur les comportements d'activité (incitations à l'activité ...) ou sur la démographie (natalité, migrations)
- Le non-recours

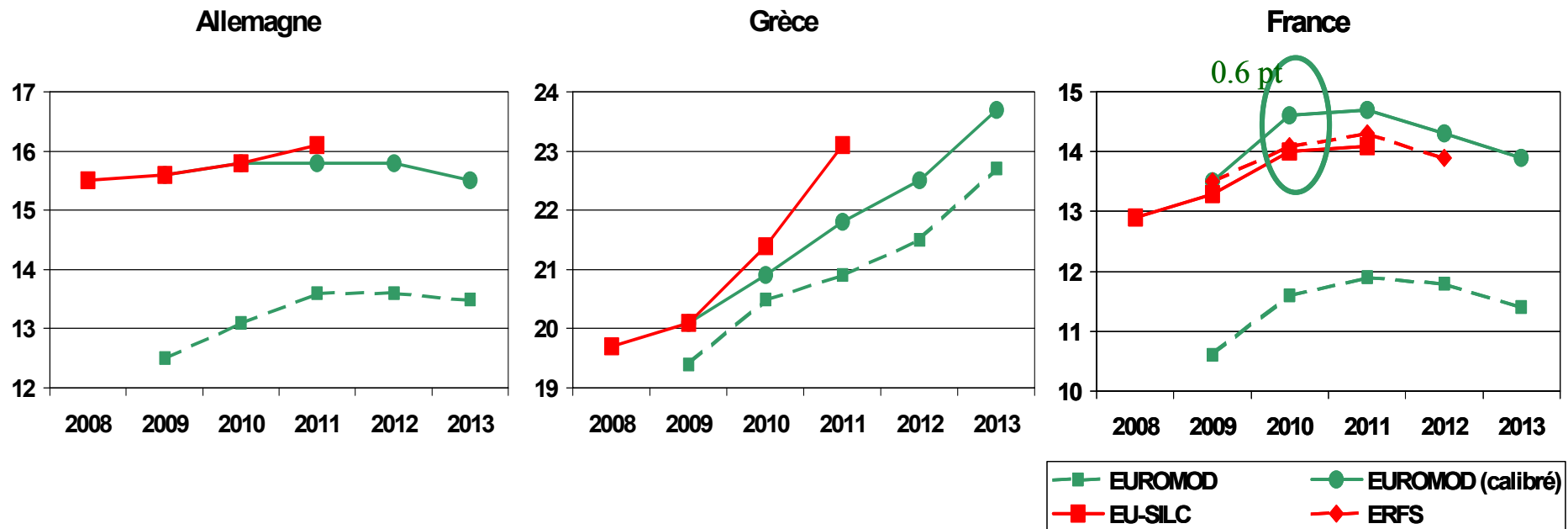
Nowcasting avec EUROMOD (résultats de décembre 2013)

- Modèle de l'Union Européenne
- Échantillon représentatif = EU-SILC 2010 (revenus 2009)
- Simulation pour 13 pays sur 2010-2013
- Pas de modification de la structure démographique
- Mise à jour dynamique de la structure de l'emploi : transitions aléatoires au niveau individuel en fonction du sexe, de l'âge et du niveau d'éducation entre chômage et emploi (modèle estimé sur EU-SILC)
- Données conjoncturelles pour mettre à jour les revenus non-simulés
- Calibrage au niveau de chaque ménage pour que le niveau de vie simulé en 2009 corresponde à celui de EU-SILC 2009

- Résultats : taux de pauvreté, niveaux de vie médians et moyens

Nowcasting avec EUROMOD (résultats de décembre 2013)

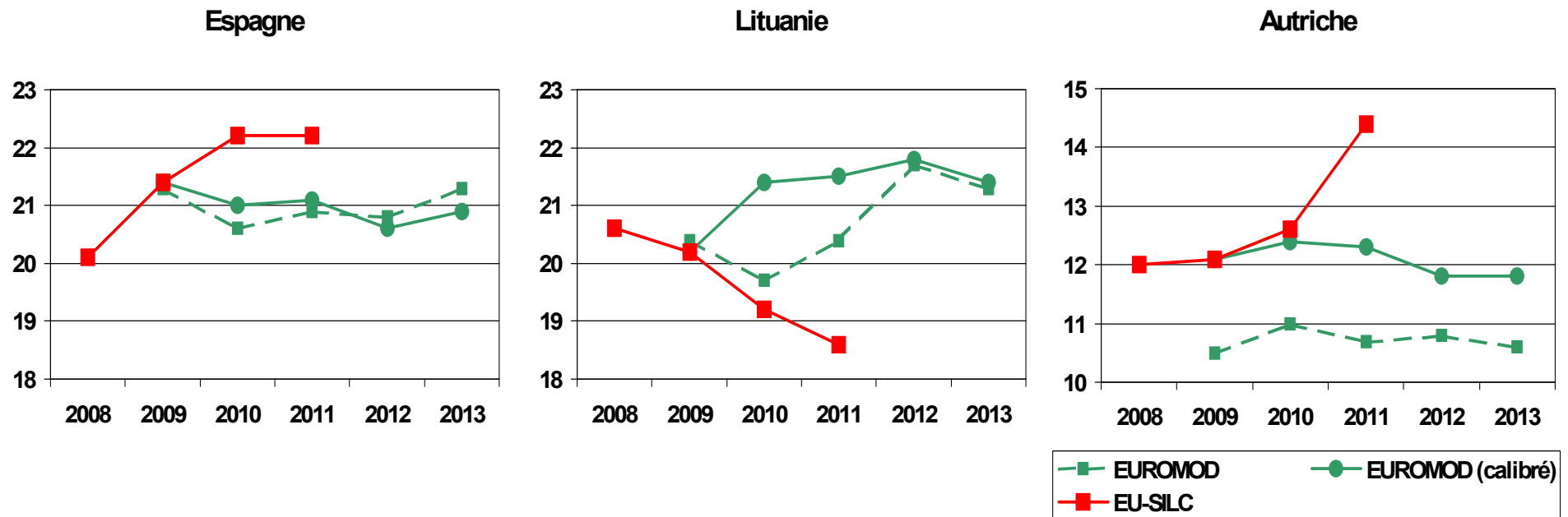
Taux de pauvreté



Nowcasting: estimating developments in the risk of poverty and income distribution in 2012 and 2013, Research note 1/2013, C. Leventi, J. Navicke, O. Rastigina, H. Sutherland (decembre 2013).

Nowcasting avec EUROMOD (résultats de décembre 2013)

Taux de pauvreté



Nowcasting: estimating developments in the risk of poverty and income distribution in 2012 and 2013, Research note 1/2013, C. Leventi, J. Navicke, O. Rastigina, H. Sutherland (decembre 2013).

Nowcasting avec **EUROMOD** (résultats de décembre 2013)

Forces

- Prévisions de bonne qualité pour certains pays, y compris en situation de crise économique et à horizon de 4 ans
- Comparabilité entre les 28 Etats membres

Limites

- Un « petit » échantillon (SILC = 11 000 ménages en France)
- Difficultés à simuler finement les dispositifs (en France : 3 fois trop de RSA, 13 % de trop d'impôt sur le revenu, presque 2 fois trop d'allocations chômage)
- Le calibrage permet par construction d'obtenir un taux de pauvreté vraisemblable (sinon à 10%)

Nowcasting avec Ines (Insee) : en cours, document de travail pour le premier trimestre 2015

- Modèle français, co-géré par l'Insee et la Drees
- Échantillon = ERFS
- Etape 1 : comparaison entre la microsimulation et l'appariement
 - [Ines N à partir ERFS N] VS [ERFS N]
 - Contrainte : se rendre comparable avec l'ERFS (quelques différences conceptuelles)
- Etape 2 : *nowcasting* : projection VS évolution constatée
 - [Ines N+1 ou N+2 à partir ERFS N] VS [ERFS N+1 ou N+2]
 - Écart projection VS évolution constatée =
 - écart microsimulation/appariement
 - + écart d'évolution des revenus et de la structure de la population
- Résultats : taux de pauvreté global et sur des sous-populations (par statut d'activité, tranche d'âge, sexe), distribution des niveaux de vie (quantiles)

Nowcasting avec Ines (Insee) : en cours, document de travail pour le premier trimestre 2015

Forces

- Un grand échantillon (ERFS = 50 000 ménages)
- Construit sur l'ERFS = référence française pour la production du taux de pauvreté
- Simulation fine de la législation (modèle utilisé régulièrement pour des chiffrages *ex-ante*)
- Contrôle de la qualité des simulations (même si difficultés pour le RSA et l'AAH)
- Vieillesse et mise à jour des revenus qui tiennent compte de manière statique de l'emploi, démographie, évolution des revenus non-simulés

Limites

- Celles de la microsimulation
- Résultats à paraître

Les déclarations pré-remplies: un registre provisoire

3. Registres provisoires

appariement des données d'enquête avec déclarations fiscales
pré-remplies

un projet en soi

- qualité des déclarations pré-remplies ?
- Complétude des composantes de revenus
 - Salaires, pensions, chômage : ok
- Revenus d'indépendants : ?
 - Revenus fonciers : ?
 - Revenus mobiliers : ?
 - Prestations sociales: par microsimulation ?

Duplication des ressources pour construire des données
provisoires puis des données définitives

Les déclarations pré-remplies: peuvent-elles améliorer la microsimulation ?

4. Une stratégie moins onéreuse de calage sur données externes

mix entre microsimulation et calage pour prédire la distribution en N

calage sur des marges N grâce à des sources externes

quelles marges externes ?

INES

- masses de revenu (évolution moyenne) :
 - Comptes Nationaux annuels ou trimestriels
- statut d'activité :
 - enquête emploi ? Pôle emploi ? Nécessité d'avoir un concept comparable à celui de l'enquête (ou adapter le questionnaire)
- distributions de (composantes) de revenus :
 - Salaires des DADS ?
 - Déclarations pré-remplies: salaires, pensions, chômage

Conclusions

Difficile de construire les données définitives plus vite

Quelques pistes réalistes pour prédire (nowcasting) quelques indicateurs phares (taux de pauvreté global, seuil de pauvreté global, Gini d'ensemble)

Au-delà ? Des indicateurs précoces pour quoi faire exactement ? Déterminer des politiques conjoncturelles de lutte contre la pauvreté ?



- Quelle fiabilité de la prévision est requise ?
- Quel détail des indicateurs est nécessaire ?

Est-ce réalisable ?
Pour quel coût ?

(Ex : Crise de 2009 : quelle différence une information précoce aurait-elle faite ?)



Calcul d'indicateurs précoces

Que peut-on faire ? Et dans quel but ?

Merci de votre attention

Contact

Cédric Houdré / Juliette Stehlé

phone : 0033+141175470

E-mail : cedric.houdre@insee.fr

Insee

18 bd Adolphe-Pinard
75675 Paris Cedex 14

www.insee.fr  

Informations statistiques :

www.insee.fr / Contacter l'Insee

09 72 72 4000

(coût d'un appel local)

du lundi au vendredi de 9h00 à 17h00