

XII^e Journées de Méthodologie Statistique de l'Insee



LE NÉCESSAIRE ÉCHANTILLONNAGE DES DONNÉES FISCALES :

Le cas de l'impôt sur le revenu

L. Aeberhardt, L. Guenchi, R. Lardellier

SSM des Finances Publiques - DGFiP





Microsimuler sous contraintes:

- Évaluation de mesures fiscales et prévision des recettes...
...dans des délais contraints
- Méthode : microsimulation avec un programme de taxation de l'IR
- Une base de données administratives : le POTE
 - 36 millions de foyers
 - 3 000 variables
- Six « émissions » pour constituer le POTE



Le LOU

un échantillon performant mais qui « date » de 1997



Une stratification caduque

234 strates reposant sur trois critères :

1. Foyer imposable ou non imposable
⇒ Notion aux contours flous
2. Tranche de revenu imposable augmenté des plus-values imposées à taux proportionnel
⇒ Seuils inchangés depuis 1997
3. Catégorie de revenu dominant (salaires, pensions...)
⇒ Quelle pertinence ?



Allocation de Bankier et tirage systématique

Exhaustivité pour les foyers atypiques et fortement contributeurs à l'impôt sur le revenu

+

Allocation de Bankier : allocation de Neyman avec objectifs locaux

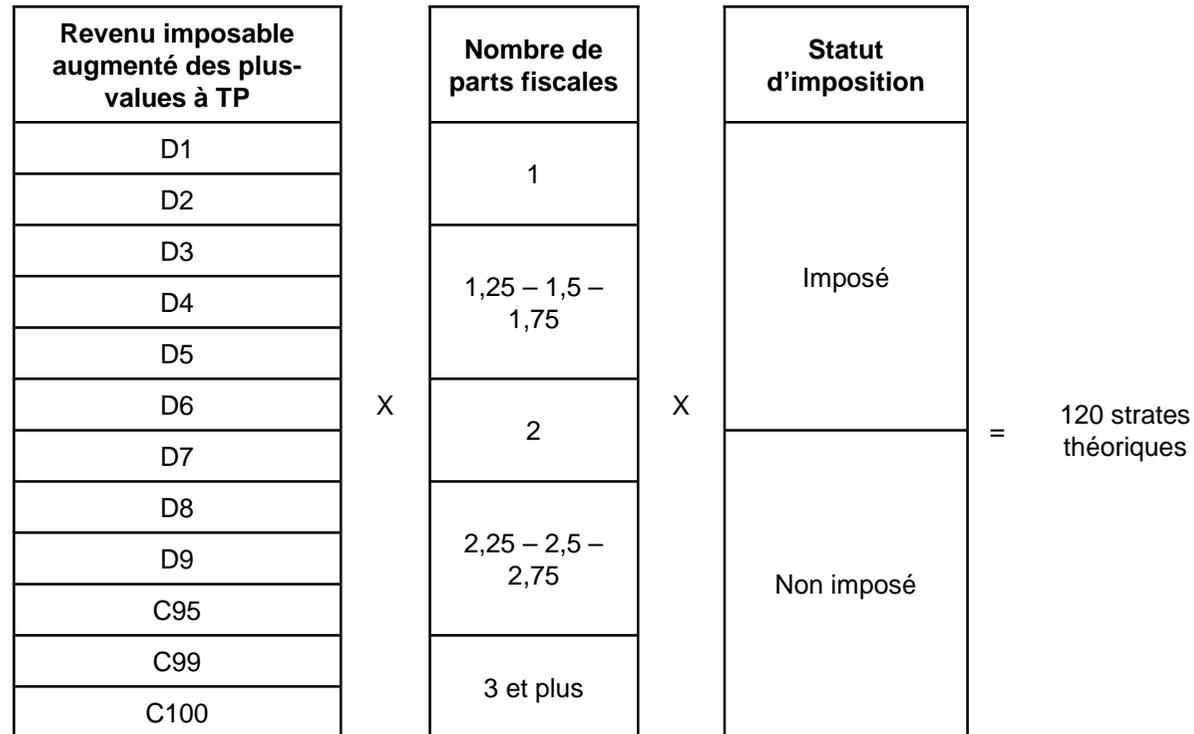
- Prise en compte des charges
 - ⇒ Problèmes d'automatisation
- Tri préalable
 - ⇒ Changements de législation
- Tirage systématique
 - ⇒ Existence de nouvelles méthodes de tirage



FÉLIN : la méthode révisée

Fichier Échantillonné de L'Impôt sur le revenu

Actualiser et automatiser la stratification



- Réduction du nombre de strates : 110 strates effectives
- Amélioration de la part de la variance intra-strates : ++ pour l'IR, + pour le RIMP, - pour le RFR
- Automatisation des seuils
- Prise en compte d'une nouvelle variable déterminante : le nombre de parts

Une allocation simplifiée

Stratégie	Variance de l'estimateur et % d'amélioration		
	Impôt sur le revenu	Revenu imposable	Revenu fiscal de référence
<i>Ancienne stratification et allocation de Bankier fonction des charges</i>	16,5	72	584
	-84,2%	-75,0%	-54,6%
<i>Option 2 avec allocation de Neyman fonction du RFR</i>	2,6	18	265

- Moindre intérêt de l'allocation de Bankier
- Abandon des charges au profit du RFR
- Majorants des précisions réelles (SAS vs CUBE)



Pourquoi et comment équilibrer ?

- Contrôle sur un nombre plus vaste de variables
- Contrôle sur les effectifs et les montants
- Cube ou FastCube ?

Équilibrer sur les déterminants de l'IR

- Les variables :

Constante	→ effectif stable
RFR, RIMP, revenus catégoriels	→ différentes définitions des revenus
RI, CI, charges, situation familiale	→ mécanismes faisant varier l'IR à revenu donné
IR	→ variable fondamentale
Taux marginal	→ non retenu

- Processus de sélection :

- Comparaison avec le POTE de plusieurs échantillons « emboîtés »
- 10 jeux de données par échantillon
- Tous les échantillons testés sont plus représentatifs que l'ancien échantillon selon les variables précitées

L'échantillon équilibré toujours préférable

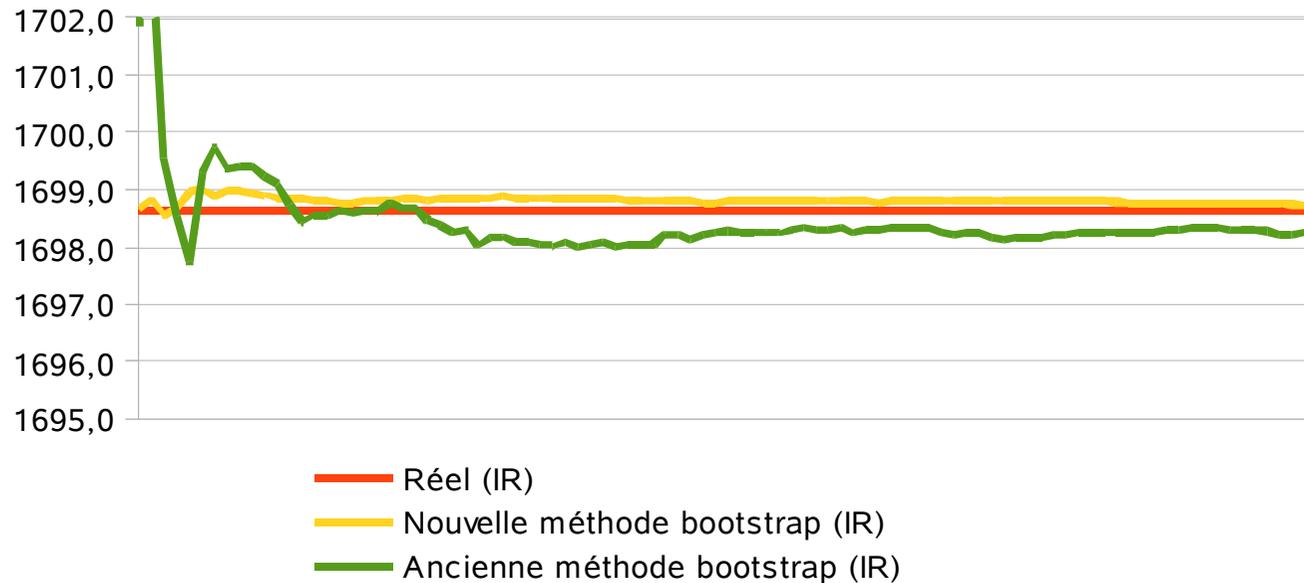
Variables d'équilibrage	Échantillons					
	AM	1	2	3	4	5
Impôt sur le revenu	5,95%	0,55%	0,52%	0,64%	0,54%	0,71%
Revenu fiscal de référence	0,91%	0,30%	0,33%	0,30%	0,41%	0,32%
Revenu imposable	0,25%	0,01%	0,01%	0,01%	0,01%	0,01%
Traitements et salaires	5,20%	1,54%	1,39%	1,39%	0,26%	0,33%
Pensions et rentes	6,73%	2,53%	2,50%	2,35%	0,19%	0,21%
Bénéfices industriels et commerciaux	18,46%	15,61%	4,11%	4,97%	5,22%	4,53%
Bénéfices non commerciaux	25,72%	25,89%	8,55%	8,62%	10,03%	9,90%
Bénéfices agricoles	33,54%	26,74%	9,21%	8,98%	8,25%	8,05%
Revenus fonciers	11,18%	10,75%	11,38%	10,60%	2,64%	2,78%
Revenus de capitaux mobiliers	9,02%	10,25%	2,99%	2,66%	2,98%	3,07%
Divers	21,00%	25,44%	27,69%	10,28%	10,28%	10,74%
Crédits d'impôt	3,98%	0,63%	0,58%	0,69%	0,57%	0,74%
Réductions d'impôt	21,08%	52,02%	46,75%	46,09%	68,71%	48,74%
Charges	6,94%	4,60%	4,68%	4,89%	4,52%	5,34%

Échantillon	Variable d'équilibrage																
	Cst	IR	RFR	RIMP	RI	CI	CH	FAM	BIC	BNC	BA	RCM	RF	DIV	TS	PR	TX
1	X	X	X	X	X	X	X	X									
2	X	X	X	X	X	X	X	X	X	X	X	X					
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X			
4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
5	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X



Validations empiriques

Estimation bootstrap des précisions



$$\text{var}_{bootstrap}(\widehat{IR}_{NM}) = 0,32$$

$$\text{var}_{bootstrap}(\widehat{IR}_{AM}) = 7,92$$

Comme attendu, les précisions estimées sont encore meilleures que les précisions théoriques anticipées

Microsimulation d'une mesure fiscale

Base de données	Gain pour l'État de la création d'une tranche à 45 % <i>(en millions d'euros)</i>
POTE 2012-5	344
LOU 2012-5	346
FÉLIN 2012-5	345

- des écarts dérisoires en apparence
- la population cible est à 97 % retenue exhaustivement dans l'échantillon
- un gain potentiel de 160 millions d'euros pour l'exercice de prévision



Perspectives

- Transformation en panel
- Vers un tirage spatialement équilibré
- Étendre le procédé à d'autres impôts: taxe d'habitation et taxe foncière