

Estimation régionale de taux de pauvreté utilisant une technique de calage

Pascal Ardilly

Insee - Département des méthodes statistiques

Objectif

Produire des estimations **régionales** annuelles de 6 indicateurs de pauvreté (demande européenne)

$$\hat{\theta} = \frac{\sum_{k \in s} w_k \cdot Y_k}{\sum_{k \in s} w_k}$$

Y_k (où k = individu) indicatrice qui détecte 4 formes de pauvreté :

- Niveau de vie inférieur à 60% du niveau de vie médian $\rightarrow \hat{\theta}_1$
- Privation matérielle « modérée » $\rightarrow \hat{\theta}_2$
- Privation matérielle « aiguë » $\rightarrow \hat{\theta}_3$
- L'appartenance à un ménage à faible intensité de travail (population de ménages spécifique) $\rightarrow \hat{\theta}_4$

S'ajoutent deux indicateurs synthétiques :

$\hat{\theta}_5$ = proportion d'individus subissant l'un au moins des 3 états de pauvreté (retenir la version 'aiguë' - et non la version 'modérée' - de la privation matérielle)

$\hat{\theta}_6$ = proportion d'individus subissant simultanément les 3 états de pauvreté.

$$\hat{\theta}_6 \leq \{\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4\} \leq \hat{\theta}_5$$

Source

Enquête européenne SILC :

- plan complexe, avec rotation (panels de 9 années)
- tirage dans le recensement (système Octopusse)
- 10 500 à 11 000 individus répondants chaque année

Tous les individus d'un ménage ont même poids ...et même Y_k .

(Très) faible effectif régional dans SILC - sauf peut-être en Ile-de-France :

SILC : nombre de ménages répondants.

Cas de quelques régions

REGIONS	2009	2010
Ile-de-France	1 682	1 729
Bourgogne	315	321
Nord Pas-de-Calais	723	788
Franche-Comté	265	261
Aquitaine	637	679
Limousin	157	166
Rhône-Alpes	850	907

Où est le problème ?

Erreur d'échantillonnage = Biais² + Variance

$$E(\hat{\theta} - \theta)^2 = (E\hat{\theta} - \theta)^2 + E(\hat{\theta} - E\hat{\theta})^2$$

Estimation 'traditionnelle' de $\theta_{\text{région}}$:

Biais = 0

Variance de $\hat{\theta}_{\text{région}}$: expression en $\frac{1}{n_{\text{région}}}$

$n_{\text{région}}$ (trop) petit \Rightarrow erreur (trop) grande

On va donc utiliser une **estimation par MODELE** pour contourner le problème, en **fondant l'estimation sur un échantillon national**.

Biais \neq 0

Variance de $\hat{\theta}_{\text{région}}$: expression en $\frac{1}{n_{\text{National}}}$

On espère ainsi réduire l'erreur d'échantillonnage !!!

Année 2010 - Taux de pauvreté monétaire

REG	Méthode directe	Source exhaustive RDL
11	10.73	12.50
21	14.88	14.61
22	20.09	14.43
23	13.64	12.99
24	11.15	11.78
25	9.40	13.31
26	13.86	12.52
31	18.65	18.55
41	16.99	13.88
42	12.80	11.29
43	13.94	12.85
52	9.12	11.17
53	13.53	11.17
54	14.53	13.83
72	12.86	12.93
73	14.78	13.97
74	18.40	14.66
82	9.54	11.82
83	14.05	14.00
91	18.15	18.63
93	14.56	15.75
94	28.35	19.32

Méthode retenue

Modèle = hypothèse \Rightarrow erreur



La relation entre la pauvreté et un ensemble de variables explicatives ne dépend pas de la région.

Formalisation :

k : identifiant de ménage

Y_k = indicatrice de pauvreté collectée par SILC

$X_k \in R^p$ = variables auxiliaires « explicatives » de Y_k ,
dont on connaît les vrais totaux régionaux
grâce à des sources externes (pseudo)
exhaustives (inclut toujours la constante).

On peut TOUJOURS définir B_{NAT} tel que

$$Y_k = B_{NAT}^t \cdot X_k + \varepsilon_{k,NAT}$$

avec $\sum_{k \in U} \varepsilon_{k,NAT} = 0$ et $\sum_{k \in U} \varepsilon_{k,NAT}^2$ minimum

où $U = U_{NAT}$ = population nationale des ménages.

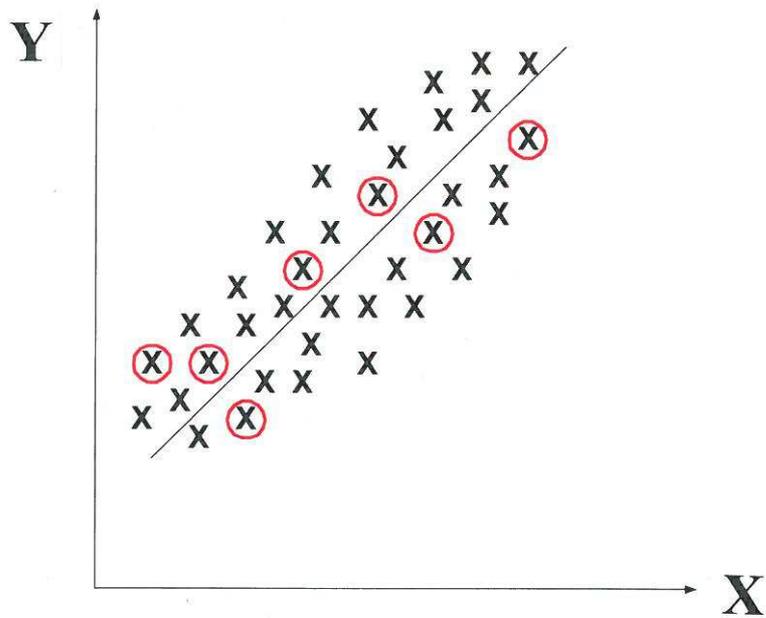
Idem pour définir B_{REG} lorsque $U = U_{REG}$ = population régionale de ménages.

Noter que Y_k peut prendre les valeurs 0 ou 1 (variables d'intérêt qualitatives).

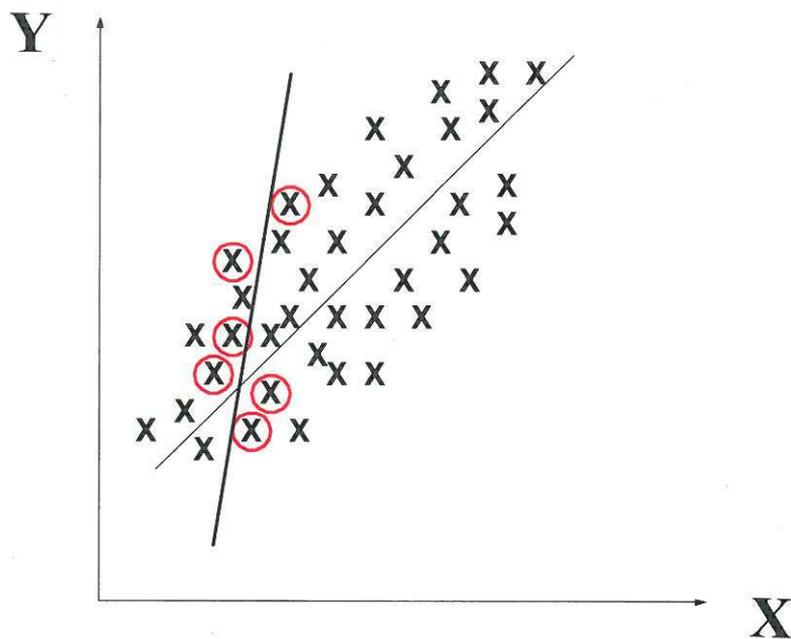
Le modèle se traduit par l'hypothèse suivante :

Cette relation s'applique à l'identique dans chaque région, c'est-à-dire

- soit on pose $B_{NAT} = B_{REG}$
- soit on pose $\sum_{k \in U_{REG}} \varepsilon_{k,NAT} = 0$



Contexte 1 : cas favorable



Contexte 2 : cas défavorable

$$Y_{REG} = \sum_{k \in U_{REG}} Y_k = B_{REG}^t \cdot \sum_{k \in U_{REG}} X_k = B_{REG}^t \cdot X_{REG}$$

$$\Rightarrow Y_{REG} = B_{NAT}^t \cdot \sum_{k \in U_{REG}} X_k \quad (\text{selon le modèle})$$

avec

$$B_{NAT} = \left(\sum_{k \in U_{NAT}} X_k \cdot X_k^t \right)^{-1} \cdot \left(\sum_{k \in U_{NAT}} X_k \cdot Y_k \right)$$

$$\Rightarrow \boxed{\hat{Y}_{REG} = \hat{B}_{NAT}^t \cdot \sum_{k \in U_{REG}} X_k = \hat{B}_{NAT}^t \cdot X_{REG}}$$

où

$$\hat{B}_{NAT} = \left(\sum_{k \in s} w_k \cdot X_k \cdot X_k^t \right)^{-1} \cdot \left(\sum_{k \in s} w_k \cdot X_k \cdot Y_k \right)$$

$s = \text{échantillon NATIONAL SILC} \Rightarrow \hat{B}_{NAT}$ aura une (très) faible variance.

\hat{Y}_{REG} est appelé **estimateur « synthétique »** de Y_{REG} .

Une subtilité :

Une condition **suffisante** de validité de l'estimateur est :
« $\epsilon_{k,NAT}$ proche de zéro ».

Mais ce n'est pas **nécessaire**, on ne demande pas que le modèle soit « bon », autrement dit les $\epsilon_{k,NAT}$ peuvent être grands !!!

En effet, même avec un modèle médiocre, si la région peut être considérée comme un ensemble « aléatoire » d'individus :

$$\frac{1}{N_{REG}} \sum_{k \in U_{REG}} \epsilon_{k,NAT} \approx \frac{1}{N_{NAT}} \sum_{k \in U_{NAT}} \epsilon_{k,NAT} = 0$$

Néanmoins, il a un **BIAIS** significatif s'il existe une composante 'fortement' explicative de la pauvreté qui est absente et dont la corrélation avec les X_k dépend sensiblement de la région.

Quelle information auxiliaire retenir ?

Celle - disponible - qui explique au mieux la pauvreté !
Trois sources ont été mobilisées.

Toute variable individuelle est agrégée au niveau ménage.

a) Le recensement

- le sexe
- l'âge (6 modalités)
- le diplôme (4 modalités)
- la nationalité (5 modalités)
- la catégorie sociale - CS (11 modalités)
- l'appartenance à une ZUS (ou non)
- la tranche d'unité urbaine (3 modalités)
- le type de ménage (5 modalités)
- le statut de locataire en HLM (ou non)

b) Le fichier Revenus Disponibles Localisés (RDL)

Fichier exhaustif issu des fichiers fiscaux, enrichi par les montants des prestations (actuellement imputés).

Pour chaque région, calcul des 19 vingtiles de la distribution **individuelle** de **niveau de vie** (= quantiles de 5 % en 5 %), où

$$\text{Niveau de vie} = \frac{\text{Revenu disponible}}{\text{Nombre d'unités de consommation}}$$

Variable auxiliaire = vecteur des 20 indicatrices d'appartenance individuelle aux 20 tranches de niveau de vie délimitées par les 19 vingtiles.

Soient deux vingtiles successifs : X_{REG} associé est égal au vingtième de la population régionale en nombre d'individus physiques (tous âges confondus).

c) Le nombre de bénéficiaires de l'Allocation de Solidarité aux Personnes Agées (ASPA)

Pour chaque région, on connaît le nombre de bénéficiaires de l'ASPA résidant en ménage ordinaire.

A l'expérience, la pondération actuelle de SILC (en France) sous-estime nettement cet effectif.

A NOTER :

Il nous a hélas manqué les effectifs régionaux :

- de l'allocation pour adulte handicapé (AAH)
- du revenu de solidarité active (RSA).

Pourquoi ?

- difficulté à séparer les allocations versées à des personnes résidant en ménage ordinaire des allocations versées à des personnes résidant en communauté.
- on trouve (source CNAF) des données régionales relatives au 31 décembre, alors que SILC repère les allocataires « durant l'année ».

Un espoir : développement du projet **Filosofi**, qui devrait fournir de telles marges à l'avenir.

Mise en œuvre ... et intervention du calage.

On recherche une méthode :

- simple et (assez) facile à expliquer (pas de modèle stochastique);
- qui ne nécessite pas de traitement statistique avancé :
 - ni le calcul des \hat{B}
(→ à renouveler pour chaque Y) ;
 - ni le calcul de A_X tel que $\hat{B} = A_X \cdot Y$
(→ programmation complexe ; volumineux).
- applicable de façon autonome par tout utilisateur final qui ne dispose pas à la fois du fichier échantillon **national** et des informations auxiliaires **régionales**.

L'atout de la méthode est son extrême simplicité de mise en œuvre, accessible au non-statisticien, car l'estimateur synthétique est connu et utilisé depuis bien longtemps !

- *Bref rappel sur le calage*

Poids initiaux d_k sans biais, c'est-à-dire

$$E\left(\sum_{k \in S} d_k \cdot Y_k\right) = \sum_{k \in U} Y_k$$

On cherche des poids w_k proches de d_k :

$$\begin{aligned} \text{Min} \quad & \sum_{k \in S} D(d_k, w_k) \\ \text{sc} \quad & \sum_{k \in S} w_k \cdot X_k = \sum_{k \in U} X_k \end{aligned}$$

On sait trouver w_k en fonction de $D(\)$
(ex : %calmar).

$$D(d_k, w_k) = \frac{(w_k - d_k)^2}{d_k} : \text{méthode « linéaire »}$$

- *Démarche*

1. on considère le fichier **national** SILC ;
2. on considère une région et ses marges X_{REG} ;
3. on inclut dans X_{REG} le nombre total de ménages N_{REG} (\Leftrightarrow *Cst* dans les régresseurs) ;
4. on « normalise » les poids : si d_k = poids du ménage k dans le fichier national
$$\forall k \in s, d_k^* = d_k \cdot \frac{N_{REG}}{N} ;$$
5. PAR REGION, on cale le fichier **national** sur les marges **régionales** X_{REG} en utilisant la méthode dite « **linéaire** » ($M = 1$);
6. on récupère un jeu de poids calés, propres à la région traitée : $w_k^{calé}$.

Propriété fondamentale

Quelle que soit la variable d'intérêt Y (quantitative ou qualitative) **on retrouve l'estimateur synthétique :**

$$\boxed{\sum_{k \in s} w_k^{calé} \cdot Y_k = \hat{B}_{NAT}^t \cdot X_{REG}}$$

Cet estimateur se justifie si et seulement si Y_k est « suffisamment bien » expliquée par les X_k .

Jamais d'échec avec la méthode linéaire !!!

Tous les individus d'un même ménage ont même poids
⇒ unité statistique du calage = ménage.

Donc tous les X_k représentent des nombres d'individus dans le ménage k , vérifiant telle ou telle modalité.

Les marges X_{REG} représentent un nombre total d'individus au niveau régional.

Deux points désagréables - mais sans conséquence :

1. les paramètres θ utilisent des Y_k **qualitatives** ;
2. il peut y avoir beaucoup de **poids négatifs** (jusqu'à 17 % des poids).

Le calage n'est qu'un instrument pratique, **l'estimation finale est *in fine* l'estimation synthétique !**

A noter

- la proportion de poids négatifs traduit l'ampleur de la différence entre les vraies structures respectives X_{REG} et X_{NAT} ;
- une méthode de calage **non linéaire** ne conduit plus à l'estimateur synthétique : elle n'apporte rien de plus, mais on peut lui trouver une justification théorique.

De fait, **autant s'en tenir à la méthode linéaire.**

Résultats

On dispose des fichiers :

- SILC 2009 : porte sur les revenus 2008.
- SILC 2010 : porte sur les revenus 2009.

Le niveau de vie des ménages / individus a été calculé selon le concept RDL ; il subsiste néanmoins une hétérogénéité sur le traitement des prestations sociales.

Le taux de pauvreté monétaire **régional** θ_1 est ‘connu’ grâce à RDL : c’est une référence pour apprécier la pertinence de la méthode.

Au niveau national

Le calage **national** a été refait avec les nouvelles marges : il ne détruit (presque) pas le calage actuel.

Calage actuel : source EEC

Calage alternatif : sources RP + RDL + ASPA

En 2009, le calage national alternatif a produit 239 poids négatifs ; en 2010, il en a produit 72.

Rapport des poids « calage alternatif / calage actuel »

Quantile	Rapport des poids (en 2010)
Maximum	3.96
99 %	1.86
95 %	1.48
90 %	1.36
75%	1.16
Médiane	0.99
25%	0.85
10 %	0.68
5%	0.52
1%	0.12
Minimum	-1.66

*Estimations nationales des tailles de population,
selon le calage national adopté*

ANNEE	Nombre de ménages		Nombre d'individus	
	Calage actuel	Calage alternatif	Calage actuel	Calage alternatif
2009	26 997 136	26 866 278	60 665 878	60 997 867
2010	27 293 225	27 106 998	60 997 389	61 298 104

Année 2009

Indicateurs nationaux de pauvreté, en %

Indicateur de pauvreté	Calage actuel	Calage alternatif
$\hat{\theta}_1$	12,89	12,77
$\hat{\theta}_2$	13,55	13,63
$\hat{\theta}_3$	5,56	5,82
$\hat{\theta}_4$	8,36	8,04
$\hat{\theta}_5$	18,47	18,40
$\hat{\theta}_6$	1,19	1,29

Année 2010

Indicateurs nationaux de pauvreté, en %

Indicateur de pauvreté	Calage actuel	Calage alternatif
$\hat{\theta}_1$	13,28	13,78
$\hat{\theta}_2$	12,62	12,47
$\hat{\theta}_3$	5,79	5,80
$\hat{\theta}_4$	9,89	9,28
$\hat{\theta}_5$	19,17	19,45
$\hat{\theta}_6$	1,52	1,44

Année 2009

*Effectifs de personnes en situation de pauvreté
Effet d'une modification du calage national*

Pauvreté au sens de ...	Calage actuel	Calage alternatif	Evolution %
θ_1	7 820 418	7 788 426	- 0,4
θ_2	8 219 312	8 319 704	+ 1,2
θ_3	3 372 182	3 549 520	+ 5,3
θ_4	3 873 144	3 816 054	- 1,5
θ_5	11 207 632	11 229 969	+ 0,2
θ_6	721 499	788 592	+ 9,3

Année 2010

*Effectifs de personnes en situation de pauvreté
Effet d'une modification du calage national*

Pauvreté au sens de ...	Calage actuel	Calage alternatif	Evolution %
θ_1	8 098 613	8 448 001	+ 4,3
θ_2	7 700 525	7 645 298	- 0,7
θ_3	3 529 922	3 554 406	+ 0,7
θ_4	4 584 882	4 393 619	- 4,2
θ_5	11 692 708	11 920 864	+ 2,0
θ_6	924 246	881 022	- 4,7

Au niveau régional

On trouve des régions spécifiques, dont l'Île-de-France et la Corse (exemples tirés de SILC 2010) :

- en Île-de-France, on recense 7 629 agriculteurs, alors que le fichier national pondéré par d_k^* en donne 84 079 : les poids des agriculteurs du fichier SILC national vont donc s'effondrer ;
- en Corse, le fichier exhaustif ASPA compte 9 934 allocataires, mais le fichier national SILC pondéré en donne seulement 927. La pondération des allocataires ASPA du fichier national va donc considérablement augmenter.

En contrepartie :

REG	Nombre de poids NEGATIFS 2009	Nombre de poids NEGATIFS 2010
11	1632	1849
21	370	331
22	668	621
23	313	136
24	380	363
25	730	755
26	677	722
31	449	288
41	364	227
42	333	200
43	509	454
52	432	282
53	552	544
54	555	571
72	359	263
73	485	398
74	547	510
82	314	135
83	540	541
91	494	435
93	658	603
94	1926	1883

Appréciation du biais de modèle : comparaison des effectifs SILC nationaux et de la somme des effectifs régionaux.

Année 2009

Biais relatif de modèle, en pourcentage

Pauvreté au sens de θ_1	Pauvreté au sens de θ_2	Pauvreté au sens de θ_3
- 0.76	1.50	5.68

Pauvreté au sens de θ_4	Pauvreté au sens de θ_5	Pauvreté au sens de θ_6
- 1.71	- 0.05	9.45

Année 2010

Biais relatif de modèle, en pourcentage

Pauvreté au sens de θ_1	Pauvreté au sens de θ_2	Pauvreté au sens de θ_3
4,91	- 0,58	0,99

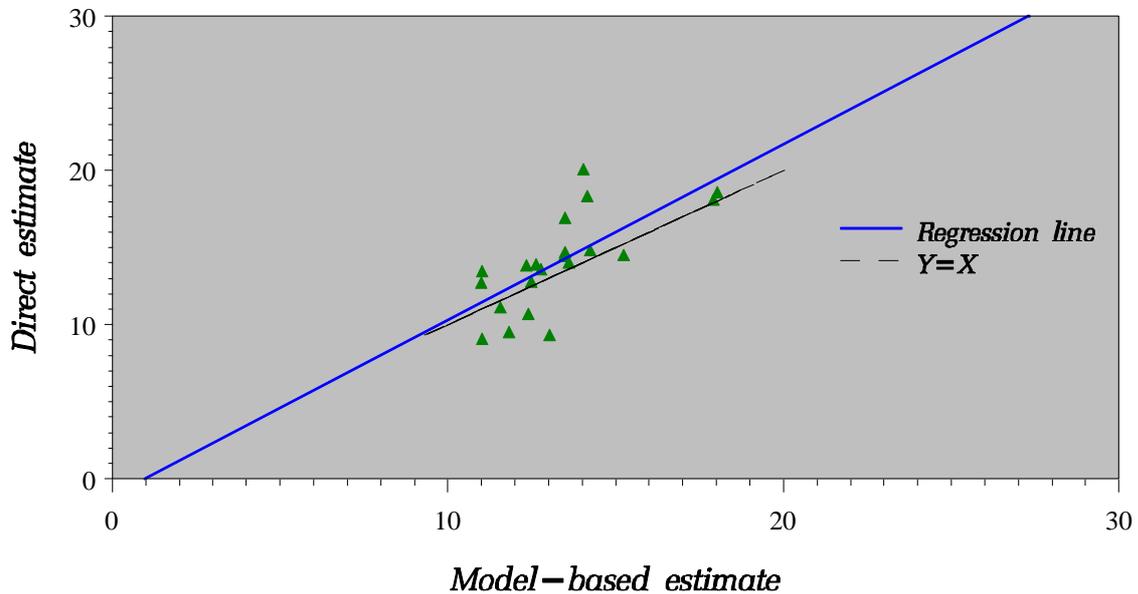
Pauvreté au sens de θ_4	Pauvreté au sens de θ_5	Pauvreté au sens de θ_6
- 4,23	2,32	- 4,16

Conclusion : les erreurs restent très acceptables - surtout celle de $\hat{\theta}_5$.

Appréciation du biais de modèle : méthode graphique.

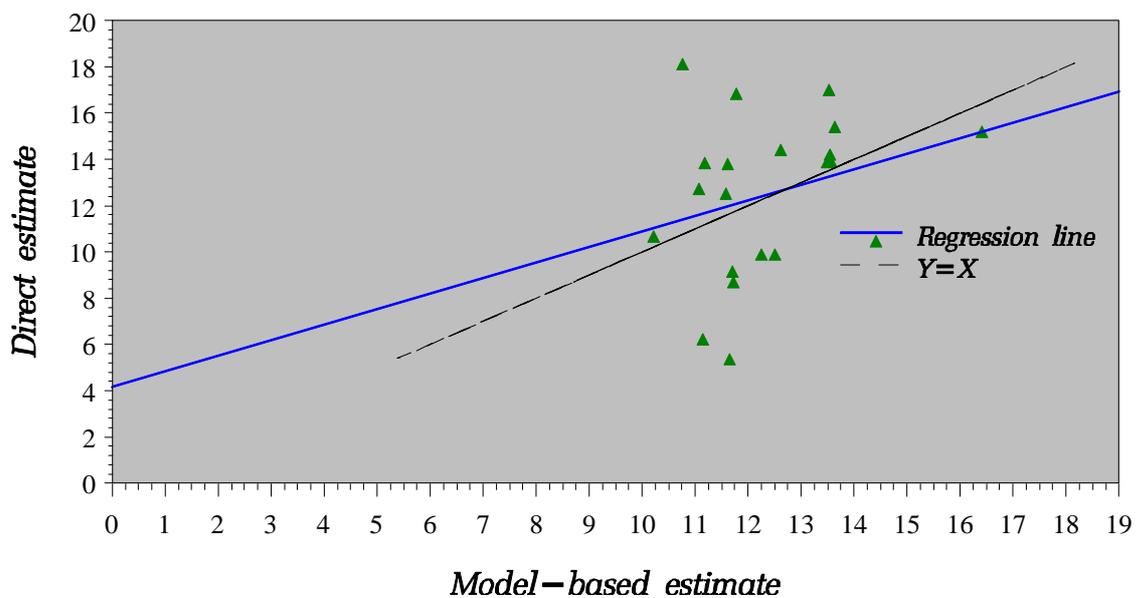
Appréciation du biais du modèle

Taux de pauvreté DIRECT versus Taux de pauvreté petits domaines
Corse absente



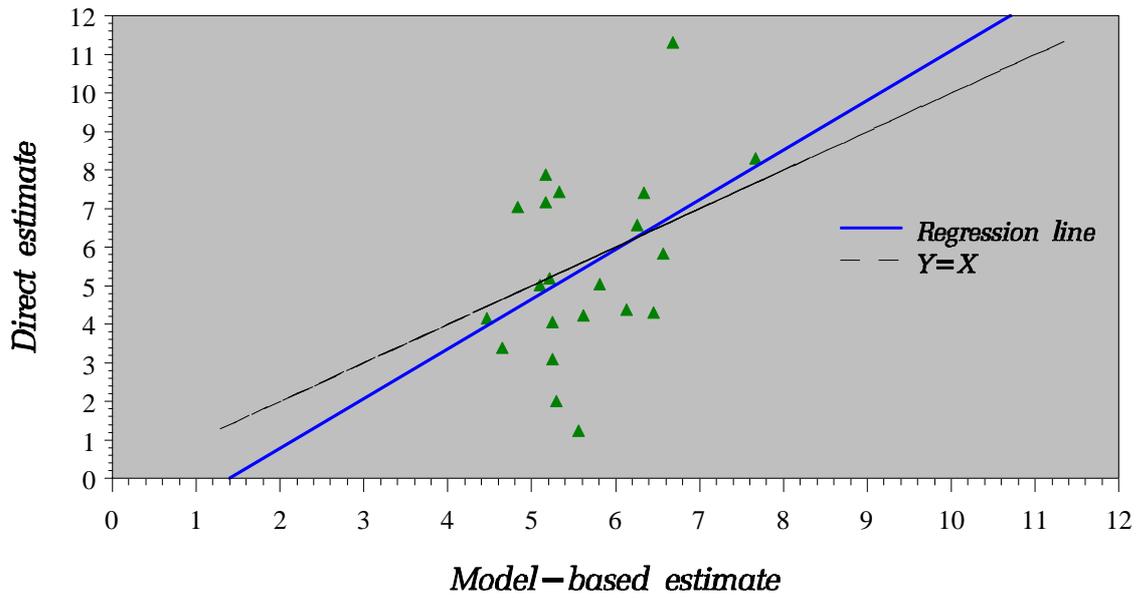
Appréciation du biais du modèle

Taux de privation modérée DIRECT versus Taux de privation modérée petits domaines
Corse absente



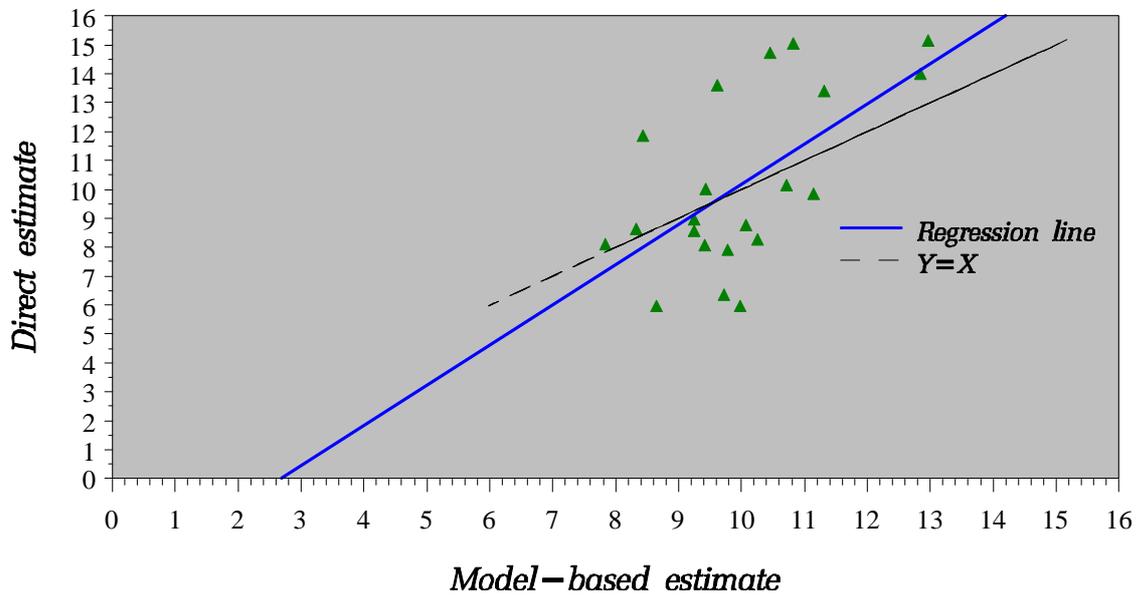
Appreciation du biais du modele

Taux de privation forte DIRECT versus Taux de privation forte petits domaines
Corse absente



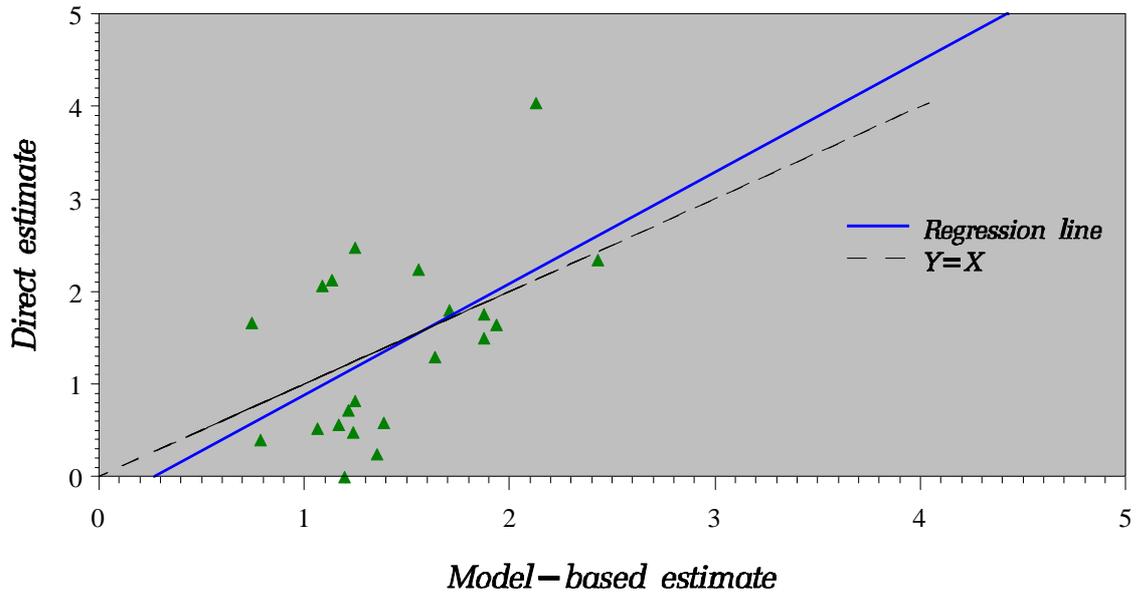
Appreciation du biais du modele

Taux de faible intensité DIRECT versus Taux de faible intensité petits domaines
Corse absente



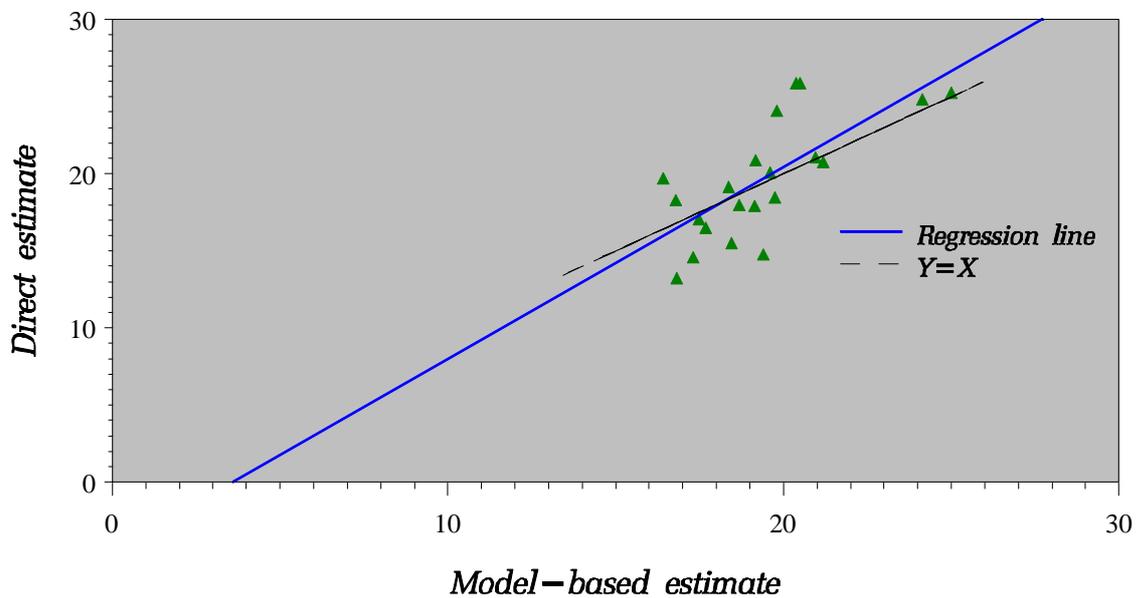
Appreciation du biais du modele

Taux de pauvreté en inter, DIRECT versus Taux de pauvreté en inter, petits domaines
Corse absente



Appreciation du biais du modele

Taux de pauvreté en UNION, DIRECT versus Taux de pauvreté en union, petits domaines
Corse absente



Mise en œuvre d'un *benchmarking* dont la cible est l'estimation nationale **avec calage alternatif** (versus calage actuel, lequel aurait redonné la statistique officielle).

$$\hat{\theta}_{REG} = \hat{\theta}_{REG} \cdot \frac{\hat{\theta}_{NAT}}{\sum_{REG} \hat{\theta}_{REG}}$$

$\hat{\theta}_{REG}$: **avant** benchmarking

$\hat{\hat{\theta}}_{REG}$: **après** benchmarking

On corrige ainsi la composante 'systématique' du biais.

Estimation des taux de pauvreté régionaux - année 2010

REG	Source RDL	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$
11	12.50	12.43	13.50	6.26	9.43	17.70	1.88
21	14.61	14.26	14.43	6.58	11.33	20.97	1.88
22	14.43	14.05	13.65	6.46	10.83	20.52	1.94
23	12.99	12.79	13.58	6.14	10.26	19.16	1.64
24	11.78	11.58	11.60	5.22	9.25	17.50	1.22
25	13.31	13.06	12.52	5.62	9.99	19.41	1.36
26	12.52	12.37	11.66	5.30	9.73	18.47	1.25
31	18.55	18.04	16.42	7.68	12.85	25.03	2.43
41	13.88	13.52	12.62	5.82	10.47	19.82	1.56
42	11.29	11.00	10.78	4.84	8.44	16.45	1.14
43	12.85	12.66	12.26	5.56	9.63	18.71	1.39
52	11.17	11.04	11.15	4.66	7.85	16.84	0.75
53	11.17	11.05	10.23	4.47	8.34	16.82	0.79
54	13.83	13.49	11.63	5.18	9.79	19.63	1.09
72	12.93	12.50	11.20	5.10	9.25	18.39	1.07
73	13.97	13.52	11.08	5.18	9.44	19.19	1.25
74	14.66	14.16	11.80	5.34	10.73	20.41	1.20
82	11.82	11.84	11.71	5.25	8.66	17.33	1.24
83	14.00	13.62	11.73	5.26	10.08	19.77	1.17
91	18.63	17.93	13.54	6.69	12.97	24.15	2.13
93	15.75	15.25	13.56	6.34	11.16	21.20	1.71
94	19.32	18.57	14.18	7.77	13.98	24.91	2.91

Les taux officiels de pauvreté RDL (cf. *insee.fr*) sont proches des taux $\hat{\theta}_1$: c'est une forme de validation de la pondération régionale. On peut penser que cette pondération réagit correctement avec les autres taux estimés $\hat{\theta}_i$.

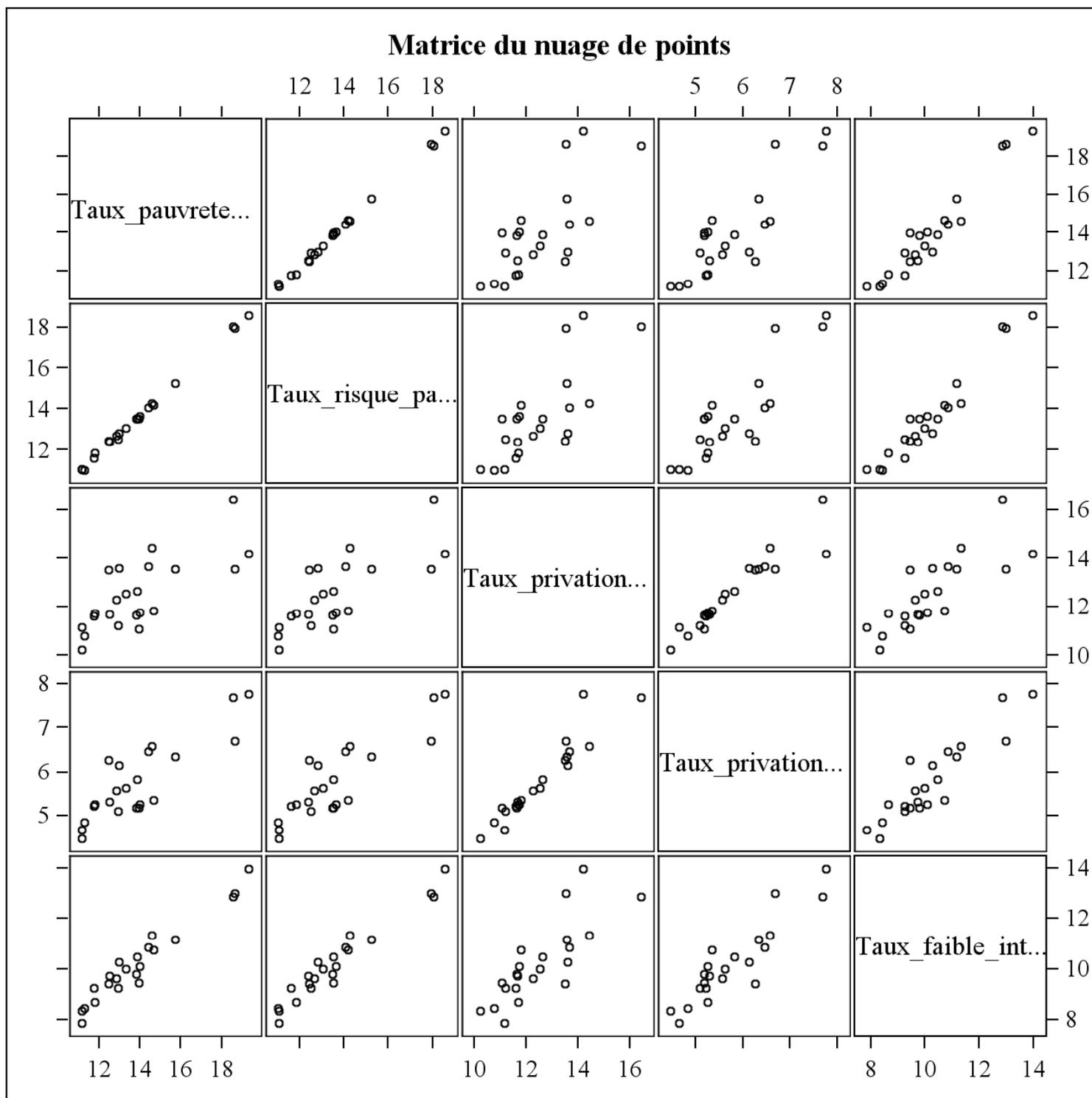
*Statistiques descriptives simples sur
les taux régionaux de pauvreté - Année 2009*

Variable	N	Moyenne	Ecart-type	Mini mum	Maxi mum
Taux de pauvreté selon RDL	22	13.41	2.5	10.60	20.00
$\hat{\theta}_1$	22	13.23	2.4	10.55	19.47
$\hat{\theta}_2$	22	13.16	1.7	10.55	17.09
$\hat{\theta}_3$	22	5.44	0.87	4.22	7.76
$\hat{\theta}_4$	22	8.54	1.5	6.42	12.73
$\hat{\theta}_5$	22	18.97	2.7	15.87	26.51
$\hat{\theta}_6$	22	1.09	0.44	0.45	2.05

*Statistiques descriptives simples sur les
taux régionaux de pauvreté - Année 2010*

Variable	N	Moyenne	Ecart -type	Mini mum	Maxi mum
Taux de pauvreté selon RDL	22	13.91	2.3	11.17	19.32
$\hat{\theta}_1$	22	13.58	2.2	11.00	18.57
$\hat{\theta}_2$	22	12.49	1.5	10.23	16.42
$\hat{\theta}_3$	22	5.76	0.89	4.47	7.77
$\hat{\theta}_4$	22	10.20	1.5	7.85	13.98
$\hat{\theta}_5$	22	19.65	2.5	16.45	25.03
$\hat{\theta}_6$	22	1.50	0.53	0.75	2.91

On croise 5 notions de taux de pauvreté : le taux de pauvreté selon RDL et les taux respectifs $\hat{\theta}_1$ à $\hat{\theta}_4$.



*Corrélations entre les différents
taux régionaux de pauvreté
Année 2010*

	Taux RDL	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$
Taux RDL	1	0.999	0.742	0.857	0.969	0.990	0.843
$\hat{\theta}_1$	0.999	1	0.761	0.870	0.970	0.991	0.854
$\hat{\theta}_2$	0.742	0.761	1	0.948	0.803	0.788	0.881
$\hat{\theta}_3$	0.857	0.870	0.948	1	0.911	0.874	0.982
$\hat{\theta}_4$	0.969	0.970	0.803	0.911	1	0.977	0.900
$\hat{\theta}_5$	0.990	0.991	0.788	0.874	0.977	1	0.844
$\hat{\theta}_6$	0.843	0.854	0.881	0.982	0.900	0.844	1

Comparer les situations régionales selon les différents critères : on attribue le rang 1 à la région la plus riche et le rang 22 à la région la plus pauvre



Rangs des régions selon la notion de pauvreté
Année 2010

REG	Rang RDL	Rang $\hat{\theta}_1$	Rang $\hat{\theta}_2$	Rang $\hat{\theta}_3$	Rang $\hat{\theta}_4$	Rang $\hat{\theta}_5$	Rang $\hat{\theta}_6$
11	6	7	15	16	7	6	17
21	17	18	21	19	19	18	18
22	16	16	19	18	17	17	19
23	10	10	18	15	14	10	15
24	4	4	6	7	6	5	8
25	11	11	13	13	12	12	12
26	7	6	8	10	10	8	10
31	20	21	22	21	20	22	21
41	13	13	14	14	15	15	14
42	3	1	2	3	3	1	5
43	8	9	12	12	9	9	13
52	1	2	4	2	1	3	1
53	2	3	1	1	2	2	2
54	12	12	7	6	11	13	4
72	9	8	5	4	5	7	3
73	14	14	3	5	8	11	11
74	18	17	11	11	16	16	7
82	5	5	9	8	4	4	9
83	15	15	10	9	13	14	6
91	21	20	16	20	21	20	20
93	19	19	17	17	18	19	16
94	22	22	20	22	22	21	22

Estimations directes
Année 2010

REG	θ_1		θ_5	
	Méthode directe	Méthode « petits domaines »	Méthode directe	Méthode « petits domaines »
11	10.73	12.43	16.51	17.70
21	14.88	14.26	21.13	20.97
22	20.09	14.05	25.95	20.52
23	13.64	12.79	17.99	19.16
24	11.15	11.58	17.12	17.50
25	9.40	13.06	14.82	19.41
26	13.86	12.37	15.55	18.47
31	18.65	18.04	25.31	25.03
41	16.99	13.52	24.12	19.82
42	12.80	11.00	19.72	16.45
43	13.94	12.66	18.04	18.71
52	9.12	11.04	13.26	16.84
53	13.53	11.05	18.31	16.82
54	14.53	13.49	20.12	19.63
72	12.86	12.50	19.17	18.39
73	14.78	13.52	20.90	19.19
74	18.40	14.16	25.93	20.41
82	9.54	11.84	14.62	17.33
83	14.05	13.62	18.50	19.77
91	18.15	17.93	24.86	24.15
93	14.56	15.25	20.82	21.20
94	28.35	18.57	41.95	24.91

En conclusion

- Une approche modèle - mais c'est incontournable ;
- Un estimateur synthétique classique;
- Une technique très simple à appliquer utilisant le calage;
- Un estimateur de (très) petite variance;
- Du biais certes, mais il reste modéré si on dispose de sources exhaustives expliquant suffisamment bien l'état de pauvreté.