Vulnérabilité énergétique des ménages

Estimation du revenu disponible des ménages à partir de modèles M-Quantiles

JMS 2015

Eric Durieux

INSEE – PSAR Analyse Territoriale

02/04/2015

Plan

- **Un contexte …. Un besoin**

- **Un exercice périlleux ?**

- **Passage à la régression quantile**

- **De la régression quantile aux M-Quantiles**

Un contexte …. Un besoin

Mesure de la vulnérabilité énergétique :

$$\text{Taux Effort Energétique} = \frac{\text{Dépenses Energétiques}}{\text{Revenu Disponible}} > \text{seuil}$$

Nécessité de modéliser chaque composante du TEE.

Un exercice périlleux?

- **Une approche carroyée non satisfaisante.**

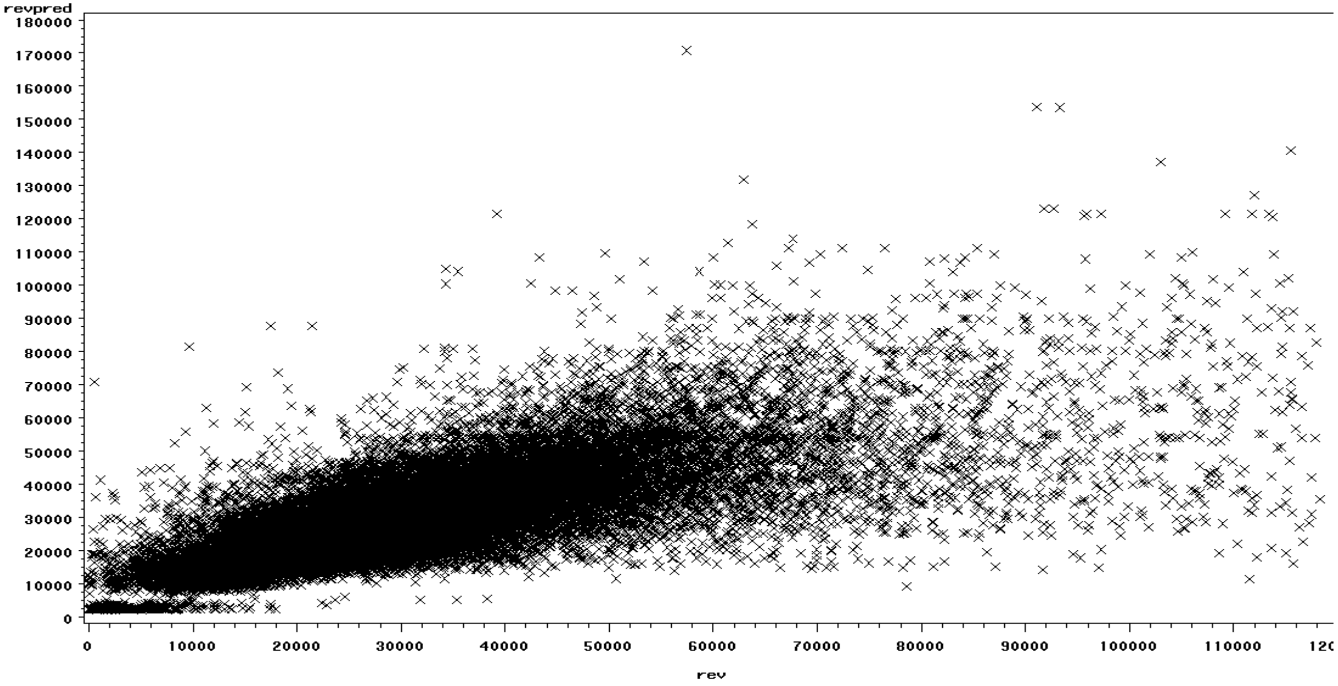
  - Non prise en compte de l'ensemble des ressources
    Revenu fiscal ≠ Revenu disponible

  - Lissage des revenus à l'échelle du carreau

- **Limites de la régression classique**

  - $$Log(\mathrm{Re}\,v) = \sum_i \alpha_i X_i$$

Un exercice périlleux?



Hétéroscédasticité des résidus

Tentative d'amélioration par segmentation du modèle

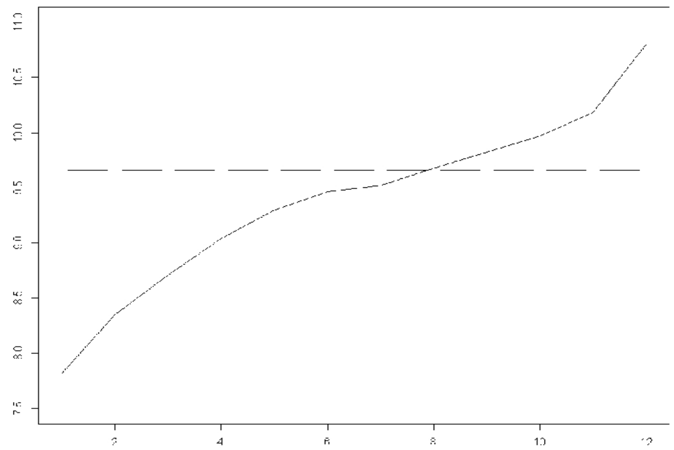
Passage à la régression quantile

➢ **Deux objectifs principaux :**

- Corriger l'hétéroscédasticité
- Sortir de la dictature de la moyenne

➢ **Rappel sur le principe :**

$$q_j(Y / X) = X\beta_j$$



Passage à la régression quantile

- **Approche du modèle en explicatif mais non prédictif**

- **Comment passer de *j* valeurs estimées correspondant au Q_j quantiles à une prédiction unique?**

- **Objet des travaux sur les M-Quantiles pour proposer des estimateurs sur petits domaines (R. Chambers & N. Tzavidis)**

De la régression quantile aux M-Quantiles

- **Travail sur des strates d'individus homogènes**

- **Un estimateurs de la variable d'intérêt peut être construit comme :**

$$\hat{m}_j = \frac{1}{N_j}\left(\sum_i y_i + \sum_i x_i \hat{\beta}_q(\hat{\theta}_j)\right)$$

$\hat{\theta}_j$ représente la moyenne des estimations issues des régressions quantiles dans la strate j pour tous les individus

- **Approche équivalente en estimant de nouveaux coefficients par fonction sur l'ensemble des coefficients (moyenne, médiane, …).**

De la régression quantile aux M-Quantiles

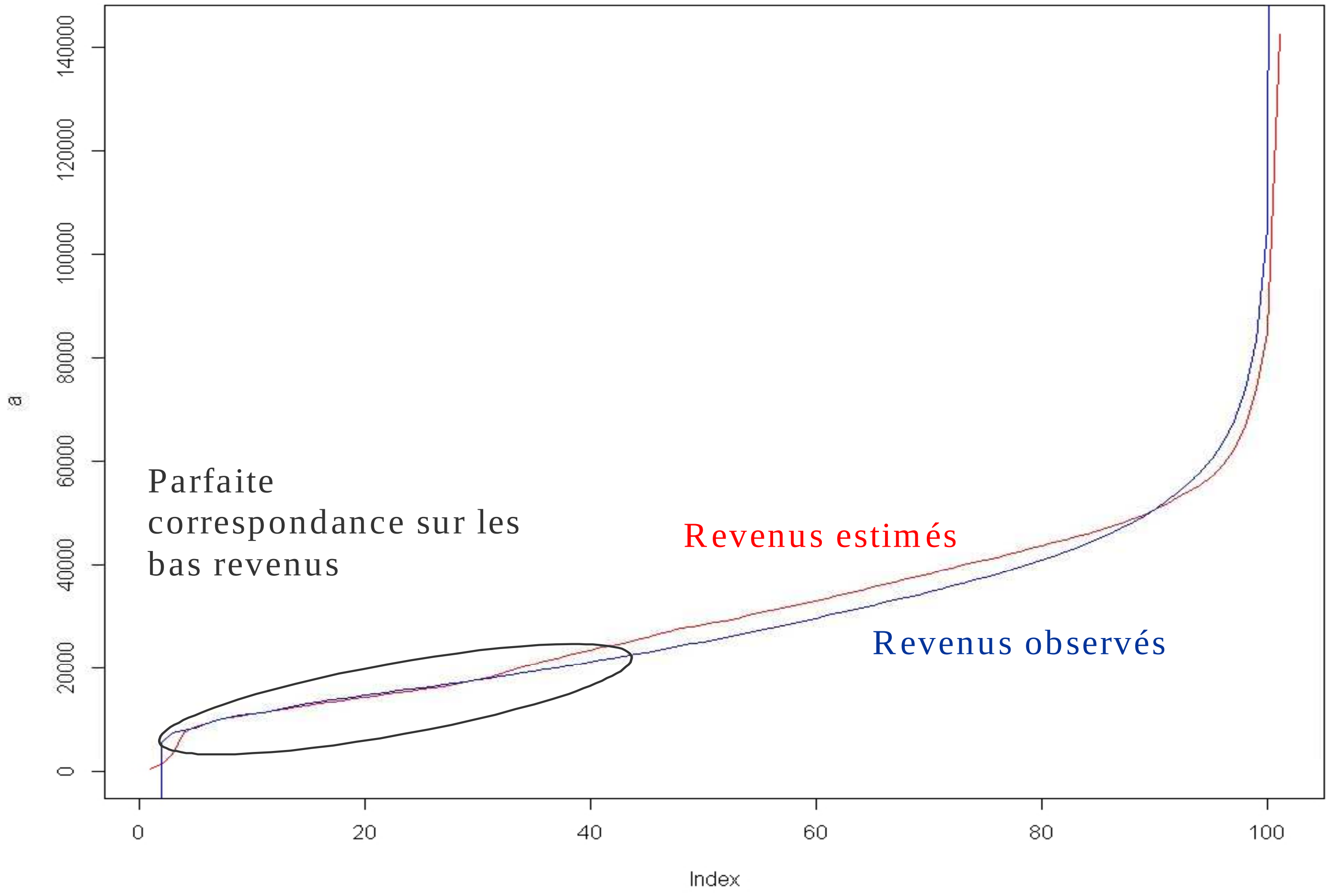
- **Choix d'une fonction sur les coefficients comme moyenne pondérée**

- **Poids de la moyenne issus d'une régression multinomiale :**
  P_i^j=probabilité que l'individu i soit sans l'intervalle borné supérieurement par Q_j

- **Principe de construction :**

$$\hat{Q}_j = \sum_i \alpha_i^j X_i \qquad \hat{\beta}_i = \sum_j P_i^j \alpha_i^j \qquad \widehat{Revenu} = \sum_i \hat{\beta}_i X_i$$

De la régression quantile aux M-Quantiles



De la régression quantile aux M-Quantiles

	Distribution observée	Distribution estimée – M-quantiles	Distribution estimée modèles linéaires simples par strates
5%	10 027	10 166	12 950
10%	12 392	11 855	14 909
20%	15 922	14 843	17 521
30%	19 128	18 391	20 067
40%	22 609	23 428	23 588
50%	26 345	28 333	26 672
60%	30 438	32 746	29 663
70%	34 953	38 004	33 744
80%	40 436	43 730	38 080
90%	47 983	52 227	45 796

Vérification de la robustesse du modèle par échantillonnage

Approche de modèles M-Quantiles