

Données transversales et vulnérabilité à la pauvreté

- À quoi se réfère-t-on par la *vulnérabilité à la pauvreté*? Différemment du niveau de la pauvreté observée, on s'intéresse plutôt à la probabilité d'être pauvre dans la future.
- En pratique, on peut se baser sur le niveau attendu bien-être et de la pauvreté selon les déterminants observés, comme le développement du capital humain, les actifs, l'environnement économique, etc.
- Comme on peut le déduire, l'évaluation du niveau de vulnérabilité nécessite d'étudier le lien statique et dynamique entre les facteurs observés et les niveaux futurs des dotations individuelles et de la pauvreté. A priori, il peut être utile de discuter de certains cas:
 - Si l'individu est effectivement observé pauvre, mais que ces dotations doivent générer des revenus plus élevés, son risque d'être pauvre à court terme et même à long terme doit être faible. C'est le cas, par exemple, d'un l'individu qui possède un diplôme supérieur et vit dans une région à faible taux de chômage.
 - Si le niveau observé du capital humain des enfants est élevé, cela peut indiquer une réduction potentielle de la pauvreté pour la prochaine génération. Ici, l'impact sur la pauvreté est à plus long terme.
- Même si l'évaluation de la vulnérabilité requiert des données de type panel, avec la rareté de ce type de données et l'abondance de données transversales dans les pays en développement, de nombreux travaux empiriques ont utilisé ces données pour évaluer le la vulnérabilité a la pauvreté.

- Parmi les modèles populaires qui ont été largement utilisés avec les données de type transversal, nous trouvons celui de Chaudhuri, 2003; Chaudhuri et al., 2002. Formellement, le niveau de vulnérabilité du ménage (ou de l'individuel) i au temps t est défini comme suit:

$$VEP_i^t = Pr(c_i^{t+1} < z).$$

- To predict the consumption of household i at time $t+1$ and the variance of consumption δ_i^2 , the following heteroscedasticity regression can be specified:
- Pour prédire la valeur prédite de la consommation du ménage i au temps $t+1$ et la variance de la consommation δ_i^2 , la régression hétéroscédastique peut être spécifiée comme suit:

$$\begin{aligned} \ln(c_i) &= X_i\beta + \varepsilon_i \\ \ln(\delta_i^2) &= X_i\theta + \zeta_i \end{aligned}$$

- En partant de ce modèle, nous pouvons observer que, avec la limitation des données, la consommation observée est simplement liée aux dotations observées. Avec l'hypothèse de normalité dans la distribution des termes d'erreur ε_i , la vulnérabilité - ou la probabilité d'être pauvre - peut être estimée par:

$$VEP_i^t = \phi\left(\frac{\ln(z) - \ln(\hat{c}_i^{t+1})}{\hat{\delta}_i^{t+1}}\right)$$

Où ϕ désigne la fonction de densité cumulée. Habituellement, nous supposons que le ménage est vulnérable à la pauvreté si son VEP dépasse la moitié.

Références:

- Baulch, B. and E. Masset (2003): "Do Monetary and Nonmonetary Indicators Tell the Same Story About Chronic Poverty? A Study of Vietnam in the 1990s," *World Development*, 31, 441–453
- Chaudhuri, S. and G. Datt (2001), "Assessing household vulnerability to poverty: a methodology and estimates for the Philippines," *Mimeo*, World Bank.
- Chaudhuri, S., J., and A. Suryahadi (2002): "Assessing Household Vulnerability to Poverty from Cross-sectional Data: A Methodology and Estimates from Indonesia," *Tech. Rep. Discussion Paper 0102-52*, Columbia University.

- En pratique, et avec l'absence de données de panel, le statut de vulnérabilité à la pauvreté de l'individu i au temps t ($V0_i^t$) est défini comme suit:

$$V0_i^t = I[VEP_i^t > 0.5] \text{ et } VEP_i^t \approx \phi\left(\frac{\ln(z) - \ln(\hat{c}_i^t)}{\hat{\delta}_i^{t+1}}\right).$$

- Enfin, à l'instar de l'indice de numérique de pauvreté, l'indice de vulnérabilité est simplement: $V0^t = E[V0_i^t]$.
- Quelles peuvent être les limites de l'approche de Chaudhuri (2003)? Comme on peut le constater, lorsque le modèle prédit parfaitement la consommation observée et que la distribution de la consommation est proche du logarithme normal, l'indice $V0$ converge vers l'indice numérique de la pauvreté.
- Une autre approche économétrique semblable à celle de Chaudhuri 2003 sera d'abord d'estimer la probabilité d'être pauvre en utilisant un modèle Probit, puis d'utiliser l'indicateur $V0_i^t = I[\Pr(poor_t^i) > 0.5]$. Nous désignons l'indice de vulnérabilité fondé sur l'approche économétrique de Chaudhuri, 2003, par $V0_{m1}^t$, et celui basé sur le modèle Probit par $V0_{m2}^t$. Dans l'application, nous montrerons comment les deux indices sont convergents dans le cas de pouvoir prédictif élevé des modèles économétriques.

- Dans un premier exercice, nous utilisons séparément chacune des deux bases de données mauritaniennes (2008 et 2014), et ce, pour l'estimation économétrique et pour évaluer le niveau de vulnérabilité. La liste des variables explicatives est la suivante (do file: Vulnerability_A1.do).

- Âge du chef de ménage
- Âge au carré du chef de ménage
- Taille du ménage
- Taille du ménage au carrée
- Nombre d'enfants de 0 à 4 ans
- Nombre d'enfants de 5 à 9 ans
- Nombre d'enfants de 10 à 17 ans
- Somme totale des échelles équivalentes adultes (échelles FAO)
- Population entre 15-64 ans
- Population entre 0 et 14 ans
- Niveau de scolarité le plus élevé atteint
- Région
- Situation familiale du chef de ménage
- Nombre de femmes adultes
- Nombre de mâles adultes

Tableau 1: Résultats avec des estimations de données transversales séparées

	2008		2014	
	Indice	Pouvoir prédictif	Indice	Pouvoir prédictif
Taux de pauvreté	44.52%		32.97%	
Vulnérabilité - indice (m1)	44.49%	$R^2=0.5086$	22.22%	$R^2=0.4348$
Vulnérabilité - indice (m2)	44.08%	$Pseudo R^2= 0.3124$	23.35%	$Pseudo R^2= 0.2573$

- Sur la base de cette première application, nous pouvons dire que le niveau de vulnérabilité a diminué en heures supplémentaires.

- Maintenant, dans le deuxième exercice, nous adoptons une hypothèse forte, dans laquelle nous supposons que la relation entre les dotations et la consommation n'a pas changé dans le temps. Formellement, nous proposons dans le deuxième exercice de considérer les probabilités prédites basées sur l'estimation avec les données regroupées de 2008 et 2014. (do file: Vulnerability_A2.do).

Table 2: Results with pooled cross section data estimation

	2008		2014	
	Indice	Pouvoir prédictif	Indice	Pouvoir prédictif
Taux de pauvreté	44.52%		32.97%	
Vulnérabilité - indice (m1)	29.33%	$R^2=0.4059$	32.72%	$R^2=0.4059$
Vulnérabilité - indice (m2)	30.63%	$Pseudo R^2= 0.2428$	33.12%	$Pseudo R^2= 0.2428$

- À première vue, il semble que ce modèle montre une augmentation de la vulnérabilité. Cependant, il faut être prudent pour les raisons suivantes:
 - En utilisant les données regroupées, on réduit les coefficients ou le rendement des dotations relativement à ce qui est observé 2014. Ceci réduit à son tour la consommation prévue, et par conséquent, augmente la pauvreté attendue en 2014.
 - Si le lien entre les variables explicatives et les changements de consommation change dans le temps (voir partie I di tableau précédent), les résultats estimés avec les données regroupées deviennent discutables.
- Maintenant, nous proposons de faire un troisième exercice dans lequel nous serons basés sur le pseudopanel construit. Rappelez-vous qu'avec le pseudopanel, on suppose que nous disposons des caractéristiques et du niveau de consommation des ménages pour chaque période. Les étapes de cette application sont (do file: Vulnerability_A3.do). ::
 1. Estimer la probabilité: $Pr\left((\hat{c}_i^{t+1}|X_i^t) < z\right)$ and $t=2008, t+1=2014$.
 2. Sur la base des coefficients estimés du modèle à l'étape 1, estimer la probabilité d'être pauvre à l'avenir après 2014: $Pr\left((\hat{c}_i^{t+2}|X_i^{t+1}) < z\right)$.

Tableau 3: Résultats avec l'estimation des données de la section transversale

	2014	
	Indice	Pouvoir prédictif
Taux de pauvreté	32.97%	
Vulnérabilité - indice (m1)	24.23%	$R^2=0.3335$
Vulnérabilité - indice (m2)	21.67%	Pseudo $R^2= 0.1871$

- Comme on peut le constater, les résultats sont proches de ceux du tableau 1 et cela conforte l'approche d'utiliser avec des données de sections distinctes pour l'estimation des modèles économétriques.