

# Conférence spéciale

## ***ENDOGENÉITE DANS UN SYSTEME D'ÉQUATIONS BIVARIÉ AVEC VARIABLES QUALITATIVES.***

*Stéfan LOLLIVIER* (\*)

(\*) ENSAE

L'endogénéité des variables explicatives pose problème lors des estimations puisque l'espérance du terme d'erreur conditionnelle à la variable explicative n'est plus nulle et les estimateurs habituels présentent des biais. Cependant, le traitement de l'endogénéité des variables explicatives est maintenant bien établi lorsque la variable expliquée et la variable explicative sont continues (doubles moindres carrés, régressions augmentées...). On dispose de procédures efficaces pour traiter ces biais, la difficulté dans la pratique consistant seulement à trouver des instruments adéquats.

Cependant, de plus en plus de travaux appliqués ont recours à des estimations économétriques faisant appel à des variables qualitatives. Celles-ci peuvent apparaître en tant que variables explicatives ou comme variables expliquées. La résolution du modèle simultané est alors beaucoup plus délicate. On dispose bien sûr de la méthode du maximum de vraisemblance, qui fournit des estimateurs asymptotiquement convergents et permet de tester l'endogénéité des variables. Mais cette technique suppose en général dans la pratique d'avoir recours à une programmation spécifique.

La difficulté peut être contournée dans un certain nombre de cas. Heckman (1978) a notamment présenté une formalisation générale permettant de traiter le problème d'endogénéité lorsque la variable expliquée est continue et la variable explicative dichotomique. L'estimateur des moindres carrés ordinaires portant sur la partie quantitative transformée est alors convergent (Heckman (1974)). La complication pratique consiste seulement à réaliser de façon rigoureuse le test permettant de conclure ou non à l'endogénéité, la matrice de variance-covariance des estimateurs n'étant pas celle qui ressort de l'estimation des moindres carrés ordinaires, en raison de l'hétéroscédasticité du terme d'erreur présent dans cette régression augmentée.

La complexité s'accroît encore d'un degré lorsque la variable expliquée est elle aussi qualitative. Même si les développements classiques restent vrais sur les variables latentes sous-jacentes aux variables dichotomiques, les problèmes pratiques liés à l'estimation se posent dès lors que l'on n'a plus affaire à des densités mais à des probabilités. Le recours à une procédure de maximisation de la vraisemblance est alors souvent incontournable.