

Modèle de régression avec variables d'écart Regression model with slack variables

Jean-Guy Loranger

Volume 50, Number 2, avril-juin 1974

URI: <https://id.erudit.org/iderudit/803042ar>

DOI: <https://doi.org/10.7202/803042ar>

[See table of contents](#)

Publisher(s)

HEC Montréal

ISSN

0001-771X (print)

1710-3991 (digital)

[Explore this journal](#)

Cite this article

Loranger, J.-G. (1974). Modèle de régression avec variables d'écart. *L'Actualité économique*, 50(2), 177–190. <https://doi.org/10.7202/803042ar>

Article abstract

A look at M. G. Dagenais' contributions (1969, 1973) on threshold regression models and at chapter 9 of S.M. Goldfeld and R.E. Quandt's book (1972) concerning switching regression models suggested to me that a new approach to estimating the threshold model by introducing slack variables might be possible. One of the main advantages of this new method is to simplify to a great extent the estimation of the likelihood function which is reduced partly to the problem of estimating a limited number of simple integrals for each iteration in the process of optimization.

In order to facilitate a better understanding of our approach, two main models will be reviewed in the next section: the twin linear probability model (which can be estimated either by OLS, by a combination of probit and OLS, or by the tobit approach) and the threshold model. A critical look at the empirical results obtained by Dagenais (1973) will also be made before closing this section.

Our new threshold model with slack variables is presented in section 3 and the main features of our new approach are summarized in the last section of this paper.

MODÈLE DE RÉGRESSION AVEC VARIABLES D'ÉCART *

1) *Introduction*

L'examen de deux contributions de M.G. Dagenais (1969, 1973) sur les modèles avec seuils de réaction et la contribution de S.M. Goldfeld et R.E. Quandt (1972, ch. 9) sur les modèles à deux ou plusieurs régimes m'ont permis de définir une approche nouvelle au modèle de régression avec seuils de réaction par une méthode que j'appellerai la méthode des écarts. L'un des principaux avantages de cette méthode est de simplifier considérablement l'estimation de la fonction de vraisemblance qui ne contient plus maintenant qu'une série d'intégrales simples à évaluer au cours de chaque itération du processus d'optimisation.

Nous examinerons d'abord, dans la prochaine section, deux principaux modèles : le modèle jumelé de régression (qui peut être estimé soit par la régression ordinaire, soit par le modèle probit et par la régression ordinaire, soit par le modèle tobit) et le modèle de M.G. Dagenais. Nous ferons également un examen critique des résultats empiriques obtenus par ce dernier modèle (Dagenais, 1973) dans le cas de l'achat d'une automobile familiale à partir de l'enquête du Michigan Survey Research Center pour 1961.

Nous présenterons ensuite, dans la section 3, notre nouveau modèle de seuil de réaction basé sur l'approche des variables d'écart pour, enfin, résumer dans la dernière section les principales caractéristiques de cette nouvelle approche.

2) *Présentation des principaux modèles de régression avec seuils de réaction*

L'idée principale à la base du modèle jumelé de régression (*twin linear probability model*)¹ et du modèle avec seuils de réaction de Dage-

* Je remercie les professeurs Marcel G. Dagenais et Richard E. Quandt pour les critiques qu'ils ont formulées à partir d'une première version de cette contribution. J'espère avoir réussi le mieux possible à prendre en considération un bon nombre de leurs remarques fort judicieuses, mais je demeure le seul responsable de toute erreur possible qui pourrait subsister et des conclusions qui apparaissent dans cet article.

1. Pour une bonne présentation sur ce sujet, le lecteur pourra se référer à Goldberger (1964), pp. 248-255.

mais est de tenter d'expliquer adéquatement le comportement discontinu d'un agent économique : soit un monopoleur qui change subitement le prix de vente de son produit, soit un consommateur qui décide d'acheter ou de ne pas acheter un bien durable.

Modèle jumelé de régression

Dans une première étape, le modèle jumelé de régression détermine la probabilité d'une action tandis que, dans une seconde étape, le modèle fournit une évaluation quantitative de la décision de ceux qui ont opté pour une action.

La première phase du modèle jumelé de régression consiste à spécifier une relation de comportement qui discrimine entre l'inaction et l'action d'agents économiques compte tenu d'un vecteur de variables explicatives X_i qui motive la décision. Soit donc :

$$y_i = X_i' \beta + u_i, \quad i \in n \quad (2.1)$$

où y_i prend les deux seules valeurs possibles : 0 (inaction) ou 1 (action), $X_i' = (X_{0i} \ X_{1i} \ X_{2i} \dots \ X_{ki})$ est le vecteur des variables explicatives et u_i est une variable aléatoire non observable, $E(u) = 0$.

L'hypothèse $E(u) = 0$ combinée avec les valeurs particulières de la variable aléatoire u_i nous donne la distribution de probabilité suivante :

	u_i	$f(u_i)$
$i \in n_1$	$-X_i' \beta$	$1 - X_i' \beta$
$i \in n_2$	$1 - X_i' \beta$	$X_i' \beta$

La distribution de probabilité de u_i implique que les variances sont hétéroscédastiques. En effet :

$$\begin{aligned} \text{Var}(u_i) &= E(u_i^2) = u_i^2 f(u_i) + u_i^2 f(u_i) \\ &\quad i \in n_1 \quad i \in n_2 \\ &= (-X_i' \beta)^2 (1 - X_i' \beta) + (1 - X_i' \beta)^2 X_i' \beta \\ &= X_i' \beta (1 - X_i' \beta), \quad \forall i. \end{aligned} \quad (2.2)$$

D'autre part, étant donné que X_i ($\forall i \in n$) est défini comme un vecteur non stochastique, l'espérance conditionnelle de y_i étant donné X_i est interprétée comme une probabilité sur y_i i.e. :

$$E(y_i | X_i) = \text{Prob}(y_i = 0 | X_i) \quad i \in n_1 \quad (2.3)$$

$$E(y_i | X_i) = \text{Prob}(y_i = 1 | X_i) \quad i \in n_2. \quad (2.4)$$

2. En effet :

$$E(u_i) = -X_i' \beta (1 - X_i' \beta) + (1 - X_i' \beta) X_i' \beta = 0 \quad \forall i.$$

En estimant le vecteur β par l'estimateur M.C. ou M.C.G. on pourrait obtenir certaines valeurs de $\hat{y}_i = X_i' \hat{\beta}$ qui se situent en dehors de l'intervalle $(0, 1)$ ce qui violerait l'hypothèse d'un modèle probabiliste de régression. Pour obvier à cette difficulté, on utilise l'approche du modèle probit en définissant une variable aléatoire continue $L \sim N(0, 1)$ telle que :

$$y_i = 0 \quad \text{si } X_i' \beta < L \quad (2.5)$$

$$y_i = 1 \quad \text{si } X_i' \beta \geq L \quad (2.6)$$

Donc :

$$\text{Prob}(y_i = 0 | X_i) = \text{Prob}(L > X_i' \beta) = 1 - \text{Prob}(L \leq X_i' \beta) \quad (2.7)$$

et

$$\text{Prob}(y_i = 1 | X_i) = \text{Prob}(L \leq X_i' \beta). \quad (2.8)$$

Mais :

$$\text{Prob}(L \leq X_i' \beta) = \int_{-\infty}^{X_i' \beta} (2\pi)^{-1/2} e^{-L^2/2} dL. \quad (2.9)$$

On peut alors estimer le vecteur β par le maximum de vraisemblance et la fonction de vraisemblance est :

$$\begin{aligned} L(y) &= \prod_{i \in n_1} \text{Prob}(y_i = 0 | X_i) \prod_{i \in n_2} \text{Prob}(y_i = 1 | X_i) \\ &= \prod_{i \in n_1} \left[1 - \int_{-\infty}^{X_i' \beta} (2\pi)^{-1/2} e^{-L^2/2} dL \right] \\ &\quad \prod_{i \in n_2} \left[\int_{-\infty}^{X_i' \beta} (2\pi)^{-1/2} e^{-L^2/2} dL \right]^3 \end{aligned} \quad (2.10)$$

L'estimation du modèle de probabilité par l'approche probit ne constitue en fait que la première étape du modèle jumelé de régression :

3. Le lecteur pourra se référer à Goldfeld et Quandt (1972) pour choisir l'une ou l'autre des méthodes de calcul numérique employées pour obtenir l'estimateur M.V. de β .

c'est l'étape qui nous permet d'établir que les $\hat{y}_i = X_i' \hat{\beta}$ sont des probabilités que $0 < y_i \leq 1$. La seconde étape du modèle jumelé de régression consiste à estimer un modèle linéaire de régression à partir d'un échantillon restreint au sous-ensemble d'observations n_2 , i.e. aux cas où $y_i > 0 \forall i \in n_2$. La valeur de la variable dépendante n'est donc plus limitée à une seule valeur (i.e. 1) mais peut prendre n'importe quelle valeur positive entre 0 et ∞ . Soit :

$$y_i^* = X_i' \beta^* + u_i \quad i \in n_2 \quad (2.11)$$

$$E(u_i) = 0 \quad (2.12)$$

$$\text{Var}(u_i) = \sigma^2. \quad (2.13)$$

Si on estime β^* par M.C., il n'y a aucune contrainte pouvant garantir une valeur positive pour $\hat{y}_i^* = X_i' \hat{\beta}^*$. Pour obvier à cette difficulté, Tobin (1958) a proposé de remplacer l'estimation du modèle jumelé de régression par l'approche suivante (modèle tobit).

Soit : L une variable aléatoire continue $N(0, \sigma^2)$

$$y_i = 0 \text{ si } X_i' \beta < L \quad i \in n_1 \quad (2.14)$$

$$y_i = X_i' \beta - L_i \text{ si } X_i' \beta \geq L \quad i \in n_2 \quad (2.15)$$

$$\begin{aligned} \text{Prob}(y_i = 0 | X_i) &= \text{Prob}(L > X_i' \beta | X_i) \\ &= 1 - \text{Prob}(L \leq X_i' \beta | X_i), \quad i \in n_1 \end{aligned} \quad (2.16)$$

$$\text{Prob}(y_i > 0 | X_i) = \text{Prob}(L = X_i' \beta - y_i | X_i), \quad i \in n_2 \quad (2.17)$$

En posant $v = (L/\sigma)$ une variable aléatoire $N(0, 1)$, la fonction de densité est : $f(v) = (2\pi)^{-1/2} e^{-v^2/2}$ et la probabilité de (L/σ) est $f(v)dv$. On notera que pour $i \in n_2$, $L/\sigma = (X_i' \beta - y_i)/\sigma$. On peut estimer le vecteur β par le maximum de vraisemblance en définissant la fonction de vraisemblance suivante :

$$\begin{aligned} L(y) &= \prod_{i \in n_1} \text{Prob}(y_i = 0 | X_i) \prod_{i \in n_2} \text{Prob}(y_i > 0 | X_i) \\ L(y) &= \prod_{i \in n_1} \left[1 - \int_{-\infty}^{X_i' \beta / \sigma} (2\pi)^{-1/2} e^{-v^2/2} dv \right] \\ &\quad \prod_{i \in n_2} \left[\left(\frac{1}{\sigma} \right) (2\pi)^{-1/2} e^{-1/2 [(X_i' \beta - y_i)/\sigma]^2} dL \right]. \end{aligned} \quad (2.18)$$

Modèle de régression avec seuils de réaction

La contribution de Marcel G. Dagenais (1969, 1973) sur les modèles avec seuils de réaction constitue sans doute une extension importante du modèle jumelé de régression ou des modèles non linéaires de type probit ou tobit.

En modifiant quelque peu la notation de Dagenais (1973) le modèle avec double seuil de réaction se présente de la manière suivante. Soient : W_i ($i \in n$) la variable dépendante aléatoire, $X_i = (X_{0i}, X_{1i}, \dots, X_{ki})$ le vecteur des variables explicatives de la régression, ($i \in n$), les X_i sont non stochastiques, $\beta' = (\beta_0, \beta_1 \dots \beta_k)$ le vecteur des coefficients de la régression, L une constante représentant une valeur limite que peut prendre y_i , u_i ($i \in n$) l'erreur de spécification de la régression, v_i ($i \in n_1 \subset n$) un seuil de réaction inférieur à la limite L tel que, pour tout $W_i < L - v_i$, $W_i = y_{1i} - u_i$, s_i ($i \in n_2 \subset n$) un seuil de réaction supérieur à la limite L tel que, pour tout $W_i > L + s_i$, $W_i = y_{2i} - u_i$. De plus, les seuils de réaction v_i et s_i sont tels que, pour tout $L - v_i \leq W_i \leq L + s_i$, ($i \in n_3 = n - n_1 - n_2$), $W_i = L$. En conséquence, on définit $y_{1i} = X_i' \beta$ et $y_{2i} = X_i' \beta^*$. Le vecteur β^* est différent de β mais la différence peut se réduire à la seule différence entre les deux termes constants⁴.

La représentation graphique de ce modèle dans un espace à deux dimensions apparaît au graphique 1 en supposant que, exception faite du terme constant, les autres composantes du vecteur X_i se réduisent à un scalaire.

On notera dans ce graphique que nous avons supposé que $L > 0$, $L + s = 2L$ et $L - v = 0$. De plus, nous avons indiqué à la gauche du graphique les sous-ensembles d'observations de W_i propres à chaque partie du graphique.

Dagenais fait l'hypothèse additionnelle que les seuils v_i et s_i sont des variables aléatoires qui peuvent être expliquées par des modèles de régression propres à chaque seuil :

$$\begin{aligned} v_i &= R_i' \xi - \eta_{1i} \geq 0 & \forall i \in n_1 \\ s_i &= Z_i' \theta + \eta_{2i} \geq 0 & \forall i \in n_2 \end{aligned}$$

et où R_i et Z_i sont des vecteurs de variables explicatives (non stochastiques) qui peuvent être soit entièrement différents de X_i ou en partie contenir quelques composantes de X_i . Soit donc $R_i' = (R_{0i}, R_{1i}, \dots, R_{ki})$ le vecteur des variables explicatives relatives au seuil inférieur v_i et

4. Dagenais se limite à différencier β et β^* par la différence dans le terme constant. Mais il n'y a aucune raison qui puisse justifier a priori cette limitation. On peut très bien concevoir une pente pour la droite de régression pour $i \in n_1$ et une autre pour $i \in n_2$ puisque, de toute manière, la pente est 0 pour $i \in n_3$.

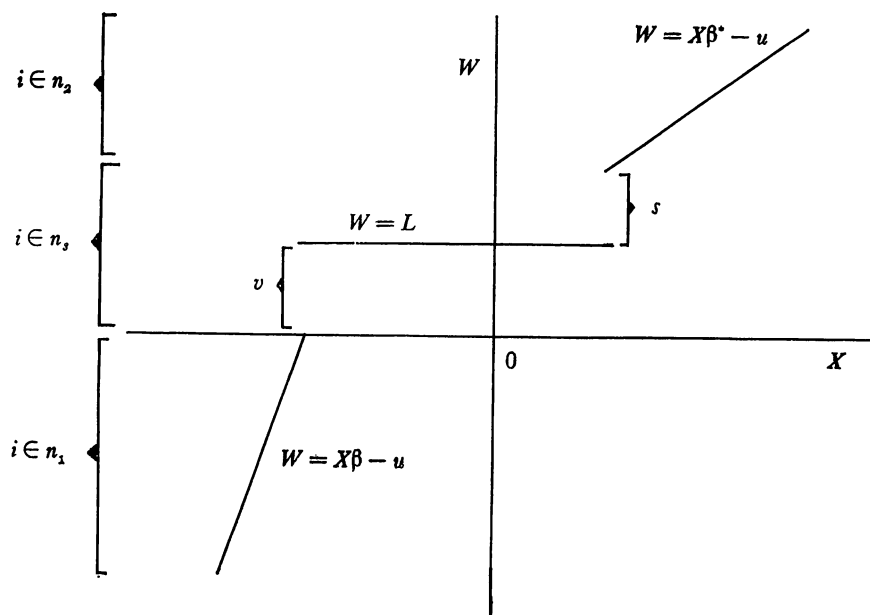
$Z'_i = (Z_{0i} \ Z_{1i} \ \dots \ Z_{mi})$ le vecteur des variables explicatives relatives au seuil supérieur s_i .

Etant donné les contraintes imposées sur η_{1i} , η_{2i} et sur u_i , Dagenais est forcé de supposer que (u, η_1, η_2) a une distribution normale conjointe tronquée⁵. Dagenais propose d'estimer le vecteur des paramètres $(\beta, \beta^*, \xi, \theta)$ et le vecteur des variances-covariances $\sigma_u^2, \sigma_{\eta_1}^2, \sigma_{\eta_2}^2, \sigma_{u\eta_1}, \sigma_{u\eta_2}, \sigma_{\eta_1\eta_2}$ par le maximum de vraisemblance. La fonction de vraisemblance décrite dans la version originale de 1969 contenait une somme de trois intégrales triples et trois intégrales doubles à évaluer. La nouvelle version de 1973 est réduite maintenant à une somme de deux intégrales doubles et trois intégrales à évaluer. C'est un progrès énorme du point de vue calcul numérique !

Critiques des résultats empiriques de Dagenais (1973)

Les résultats obtenus par Dagenais en appliquant son modèle avec un seul seuil (supérieur) à l'achat d'une automobile familiale sont plutôt décevants car, pour assurer une convergence raisonnable de la

GRAPHIQUE 1



5. Pour une présentation explicite de cette distribution tronquée, le lecteur se référera à Dagenais (1973), p. 3.

fonction de vraisemblance⁶, il lui a fallu poser à priori l'hypothèse de l'absence de corrélation entre u et η et l'hypothèse d'une variance commune aux deux variables aléatoires i.e. $\sigma_u^2 = \sigma_\eta^2$ (1973, p. 12).

L'imposition de telles restrictions à priori diminue considérablement l'originalité du modèle de Dagenais dont l'une des principales différences par rapport au modèle jumelé de régression était précisément de tenir compte de la dépendance probable entre les deux variables aléatoires. Evidemment, l'absence de corrélation n'implique pas nécessairement l'indépendance entre u et η surtout avec des distributions tronquées, mais cette hypothèse d'absence de corrélation fait peser de graves doutes dans notre esprit.

De plus, après les premières applications empiriques, Dagenais s'est vite rendu compte qu'un meilleur ajustement d'un modèle plus réaliste serait celui qui introduit la même variable explicative clef (R = revenu permanent/valeur de la voiture de la période précédente) à la fois dans le modèle expliquant l'achat et dans la fonction conditionnant le seuil, de telle sorte que les variables qui expliquent le seuil peuvent être interchangeables avec celles qui expliquent l'achat. En conséquence, si on doit retrouver en pratique à peu près les mêmes variables (avec cependant des coefficients différents) dans la fonction du seuil et dans la fonction d'achat, le modèle avec seuil de réaction de Dagenais ressemble de plus en plus au modèle jumelé de régression.

Finalement, Dagenais note qu'un modèle jumelé du type probit et de régression à partir d'une loi normale tronquée à la Cragg et Uhler (1969) produit des résultats aussi bons (et nous pourrions ajouter possiblement meilleurs) que son modèle avec seuil de réaction lorsqu'appliqué à l'achat d'une voiture familiale.

3) Définition d'un modèle de seuil de réaction par les variables d'écart

Nous allons dans les lignes qui suivent présenter une formulation différente du modèle de Dagenais en introduisant des variables d'écart

6. La fonction de vraisemblance à optimiser dans le cas d'un modèle avec un seuil (supérieur) de réaction s'écrit :

$$\varphi = \sum_{i=1}^{n_3} \ln \left\{ \int_{-\infty}^{y_i - L_i} \int_{y_i - L_i - u_i - Z_i' \theta}^{\infty} f(u, \eta) d\eta du + \int_{y_i - L_i}^{\infty} f(u) du \right\} \\ + \sum_{j=1}^{n_2} \ln \left\{ \int_{-\infty}^{W_j - L_j - Z_j' \theta} f(y_j - W_j, \eta) d\eta \right\}$$

Voir Dagenais (1973), p. 6.

qui transforment le modèle de Dagenais en un modèle à deux régimes que nous tenterons d'estimer par l'approche nouvelle de Goldfeld et Quandt (1972, ch. 9). Nous limiterons notre présentation à un modèle avec un seul seuil (supérieur) de réaction, mais il est évident que notre approche peut être étendue pour couvrir des modèles à deux seuils ou à plusieurs seuils.

Définition d'un modèle à deux régimes

Soit un échantillon n qui peut être réparti en deux sous-groupes : n_1 et n_2 . Soit n_1 le sous-groupe d'agents économiques (ex. consommateurs) qui prennent la décision d'inaction. Soit L l'évaluation de cette décision (ex. $L = 0$). Soit $n_2 = n - n_1$ le sous-groupe d'agents économiques qui prennent la décision d'action. En reprenant la notation du modèle de Dagenais à un seuil (supérieur) de réaction on a :

$$W_i = L \quad \text{pour } W_i \leq L + s_i \quad i \in n_1 \quad (3.1)$$

$$W_i = Y_i - u_i \quad \text{pour } W_i > L + s_i \quad i \in n_2 \quad (3.2)$$

$$Y_i = X_i' \beta \quad i \in n_2 \quad (3.3)$$

$$s_i = R_i + \eta_i \quad \text{pour } s_i \geq 0 \quad i \in n \quad (3.4)$$

$$R_i = Z_i' \theta \quad (3.5)$$

$$W_i \text{ est aléatoire} \quad \forall i \in n.$$

Y_i est aléatoire parce que X_i' est aléatoire $\forall i \in n_2$ mais X_i est distribuée d'une façon indépendante des autres variables aléatoires (u_i , η_i , s_i , W_i). Nous préférons cette hypothèse à l'hypothèse d'un vecteur de variables explicatives non stochastiques au nom d'un plus grand réalisme économique puisque nous n'avons aucun contrôle sur les déterminants de Y_i .

$$u_i \text{ est aléatoire} \quad \forall i \in n_2, E(u_i) = 0.$$

$$s_i \text{ est aléatoire} \quad \forall i \in n \text{ parce que } \eta_i \text{ est aléatoire } \forall i \in n.$$

$$E(\eta_i) = 0, E(\eta_i^2) = \sigma_\eta^2 \text{ et } E(\eta_i \eta_j) = 0 \quad \forall i \neq j.$$

R_i est aléatoire parce que Z_i' est aléatoire $\forall i \in n$ mais Z_i est distribuée d'une façon indépendante des autres variables aléatoires (u_i , η_i , X_i , W_i). θ et β sont des vecteurs de paramètres non aléatoires⁷.

En introduisant des variables d'écarts ε_{1i} , ε_{2i} et ε_{3i} respectivement dans (3.1), (3.2) et (3.4) pour transformer les contraintes d'inégalités en égalités, cette reformulation du modèle nous évite de recourir à la

7. Ici encore nous préférons admettre l'hypothèse d'un vecteur Z_i stochastique pour être cohérent avec l'hypothèse que nous avons déjà admise sur les X_i .

nécessité de spécifier des lois normales tronquées pour les variables aléatoires du modèle :

$$W_i = L + s_i - \varepsilon_{1i} \quad i \in n_1 \quad (3.6)$$

$$W_i = L + s_i + \varepsilon_{2i} \quad i \in n_2 \quad (3.7)$$

$$s_i = R_i + \eta_i - \varepsilon_{3i} \quad i \in n \quad (3.8)$$

$$\text{où : } \varepsilon_{1i} \sim N(\mu, \sigma_1^2) \quad (3.9)$$

$$\varepsilon_{2i} \sim N(\nu, \sigma_2^2) \quad (3.10)$$

$$\varepsilon_{3i} \sim N(\rho, \sigma_3^2). \quad (3.11)$$

En substituant (3.5) dans (3.8) et (3.8) dans (3.6) et (3.7) on obtient :

$$W_i = L + Z_i' \theta + \eta_i - \varepsilon_{1i} - \varepsilon_{3i} \quad i \in n_1 \quad (3.12)$$

$$W_i = L + Z_i' \theta + \eta_i + \varepsilon_{2i} - \varepsilon_{3i} \quad i \in n_2. \quad (3.13)$$

En prenant l'espérance de (3.12) et (3.13) :

$$E(W_i) = L + E(Z_i' \theta) - (\mu + \rho) \quad i \in n_1 \quad (3.14)$$

$$E(W_i) = L + E(Z_i' \theta) + \nu - \rho \quad i \in n_2. \quad (3.15)$$

Puisque en vertu de (3.1) $E(W_i) = L$, $\forall i \in n_1$ et en vertu de (3.2)

$E(W_i) = E(Y_i) - E(u_i) = E(X_i' \beta)$, on peut récrire (3.14) et (3.15) :

$$L = L + E(Z_i' \theta) - (\mu + \rho) \quad i \in n_1 \quad (3.16)$$

$$E(X_i' \beta) = L + E(Z_i' \theta) + \nu - \rho \quad i \in n_2 \quad (3.17)$$

On déduit de (3.16) que :

$$E(Z_i' \theta) = \mu + \rho \quad (3.18)$$

et que :

$$E(X_i' \beta) = L + \mu - \nu. \quad (3.19)$$

En conséquence, l'estimation de la somme des paramètres μ et ρ correspond à la prévision conditionnelle du seuil dans le modèle de Dagenais et la somme des paramètres $(L + \mu - \nu)$ nous donne directement la prévision conditionnelle du comportement pour le sous-groupe d'agents économiques qui ont décidé d'une action positive (ex. : le montant des dépenses effectuées par le sous-groupe d'acheteurs d'automobiles familiales).

Information exogène et fonction en escalier

Nous sommes donc en présence de deux modèles ou générateurs différents des observations sur la variable dépendante W_i . On peut alors tenter d'estimer les paramètres (θ , σ^2 , μ , ν , ρ , σ_1^2 , σ_2^2 , σ_3^2 et leur covariance) par la méthode de Goldfeld et Quandt qui consiste à avoir recours à une information exogène à (3.12) et (3.13) pour pouvoir identifier par quel modèle W_i est généré.

En examinant *ex post* les valeurs de W_i , on peut identifier directement par quel modèle ou générateur W_i a été produit. Mais si on se place d'un point de vue *ex ante*, il est impossible de déterminer le régime qui pourra générer les valeurs de W_i , $\forall i \in n$. C'est le même problème que dans le cas du modèle tobit : quand nous examinons *ex post* les observations sur la variable dépendante, nous constatons qu'il y a un sous-groupe d'observations n_1 dont les valeurs sont strictement 0 et un autre sous-groupe n_2 dont les valeurs sont strictement positives. L'intérêt de l'approche de Tobin est de déterminer, d'une part, la probabilité d'avoir ($y_i = 0 \mid X_i$) et d'autre part, la probabilité d'avoir ($y_i > 0 \mid X_i$). Le recours à de l'information externe et aléatoire remplit le même rôle dans notre modèle à deux régimes que la variable aléatoire L_i dans le modèle tobit. Puisque nous avons déjà admis que la matrice d'information X est stochastique par hypothèse, on peut alors prendre l'une des variables explicatives de X comme source d'information externe au modèle à deux régimes.

Soit donc P_i ($i \in n$) la première composante principale associée à la plus grande racine caractéristique de X . Nous choisissons la première composante principale, car, par définition, $P = Xa$ où a est le vecteur caractéristique associé à la plus grande racine caractéristique. Puisque Xa est une combinaison linéaire des vecteurs colonnes de X , il appartient à la même classe de modèle linéaire que $X\beta$. En conséquence, $P_i = X_i'a$ peut être considéré comme un indicateur valable d'information externe à notre modèle à deux régimes. Soit P_0 une valeur inconnue (à estimer) telle que $P_i \leq P_0$, W_i est généré par (3.12) tandis que si $P_i > P_0$, W_i est généré par (3.13).

Soit une fonction en escalier $D(P_i) = D_i$ telle que :

$$D_i = 0 \quad \text{si } P_i \leq P_0, \quad i \in n_1 \quad (3.20)$$

et :

$$D_i = 1 \quad \text{si } P_i > P_0, \quad i \in n_2. \quad (3.21)$$

En définissant P comme une variable continue $N(P_0, \sigma^2)$, l'approximation d'une fonction en escalier par une loi de probabilité à la Goldfeld et Quandt (1972, p. 263) est :

$$D_i = D(P_i) = \int_{-\infty}^{P_i} (2\pi)^{-1/2} (\sigma^2)^{-1/2} e^{-1/2[(P-P_0)/\sigma]^2} dP. \quad (3.22)$$

Donc :

$$D_i = \int_{-\infty}^{P_i} f(P) dP = \text{Prob}(P \leq P_i). \quad (3.23)$$

Il est intéressant de noter la similitude de l'approximation de Goldfeld et Quandt avec la transformation utilisée dans les modèles probit et tobit. Dans le probit, on définit P comme une variable aléatoire continue $N(0, 1)$ et la borne supérieure d'intégration est $X_i'\beta$. Dans le tobit, P est une variable aléatoire continue $N(0, \sigma^2)$ et la borne supérieure d'intégration pour $\text{Prob}(y_i = 0 | X_i)$ est encore $X_i'\beta$. Avec l'approximation de Goldfeld et Quandt, on définit P comme une variable aléatoire continue $N(P_0, \sigma^2)$. De plus, en définissant P_i comme une composante principale de X , $\{P_i\}$ est un vecteur issu d'une combinaison linéaire des colonnes de X et, par conséquent, $P_i = g(X_i'\beta)$. On retrouve donc, en quelque sorte, la même borne supérieure pour l'intégrale que dans les modèles probit ou tobit⁸.

Estimation d'un modèle probabiliste

En multipliant (3.12) par $(1 - D_i)$ et (3.13) par D_i et en additionnant les deux modèles, on obtient le modèle probabiliste suivant :

$$W_i = L + Z_i' \theta + \eta_i - \varepsilon_{3i} - (1 - D_i) \varepsilon_{1i} + D_i \varepsilon_{2i}, \quad i \in n. \quad (3.24)$$

En posant :

$$w_i = W_i - L, \quad i \in n \quad (3.25)$$

$$\omega_i = \eta_i - \varepsilon_{3i} - (1 - D_i) \varepsilon_{1i} + D_i \varepsilon_{2i}, \quad i \in n \quad (3.26)$$

on obtient finalement :

$$w_i = Z_i' \theta + \omega_i, \quad i \in n \quad (3.27)$$

8. Cette approche de Goldfeld et Quandt ne doit pas être confondue avec une autre approche suggérée par Quandt (1972) ou la méthode λ de Quandt. Selon cette dernière approche Quandt suppose l'existence d'une probabilité inconnue ($0 \leq \lambda \leq 1$) que la nature choisisse le modèle I comme générateur d'observations et une probabilité $(1 - \lambda)$ que les observations sur Y_i soient générées par le modèle II. Swamy et Metha (1973) viennent de démontrer que dans ces conditions il n'existe pas de valeur finie pour la fonction de vraisemblance et que les paramètres du maximum de vraisemblance ne sont pas identifiables puisque la fonction de vraisemblance s'écrit alors comme la somme de 2^n combinaisons possibles de générer la variable dépendante. L'approche de Goldfeld et Quandt repose sur la *nécessité d'identifier chaque observation* à l'un ou l'autre des deux régimes tandis que la méthode λ de Quandt se limite à la *probabilité* qu'une observation soit générée par l'un ou l'autre régime.

Nous avons déjà supposé la normalité pour ε_{1i} , ε_{2i} , ε_{3i} . Il nous faut maintenant la supposer également pour η_i si on désire que ω_i soit normalement distribué de moyenne $E(\omega_i)$ et de variance $\text{Var}(\omega_i)$, $\forall i \in n$. Donc :

$$E(\omega_i) = -\rho - (1 - D_i)\mu + D_i v, \quad \forall i \in n \quad (3.28)$$

$$\text{Var}(\omega_i) = E[\omega_i + \rho + (1 - D_i)\mu - D_i v]^2 \quad (3.29)$$

$$\begin{aligned} &= \sigma_{\eta}^2 + (1 - D_i)^2 \sigma_1^2 + D_i^2 \sigma_2^2 + \sigma_3^2 - 2(1 - D_i) \sigma_{1*} \\ &+ 2D_i \sigma_{2*} - 2(1 - D_i) D_i \sigma_{12} + 2(1 - D_i) \sigma_{13} \\ &- 2D_i \sigma_{23}, \quad \forall i \in n, \end{aligned} \quad (3.30)$$

$$\begin{aligned} \text{où :} \quad \sigma_{\eta}^2 &= \text{Var}(\eta_i) & \sigma_{1*} &= \text{covar}(\eta_i, \varepsilon_{1i}) \\ \sigma_1^2 &= \text{Var}(\varepsilon_{1i}) & \sigma_{2*} &= \text{covar}(\eta_i, \varepsilon_{2i}) \\ \sigma_2^2 &= \text{Var}(\varepsilon_{2i}) & \sigma_{3*} &= \text{covar}(\eta_i, \varepsilon_{3i}) \\ \sigma_3^2 &= \text{Var}(\varepsilon_{3i}) & \sigma_{12} &= \text{covar}(\varepsilon_{1i}, \varepsilon_{2i}) \\ & & \sigma_{13} &= \text{covar}(\varepsilon_{1i}, \varepsilon_{3i}) \\ & & \sigma_{23} &= \text{covar}(\varepsilon_{2i}, \varepsilon_{3i}). \end{aligned}$$

En supposant que Z_i' est généré par un processus stochastique indépendant de celui qui génère ω_i , $\forall i \in n$, la distribution conditionnelle de $(\omega_i | Z)$ est identique à la distribution marginale de ω_i et la fonction de vraisemblance s'écrit :

$$\begin{aligned} L(\omega) &= (2\pi)^{-n/2} \prod_{i=1}^n [\text{Var}(\omega_i)]^{-1/2} \\ &\exp - \frac{1}{2} \sum_{i=1}^n \left[\frac{[w_i - Z_i' \theta + \rho + (1 - D_i)\mu - D_i v]^2}{\text{Var}(\omega_i)} \right] \end{aligned} \quad (3.31)$$

On optimise cette fonction de vraisemblance en suivant la procédure itérative déjà décrite dans Goldfeld et Quandt (1972, pp. 264-265). Les paramètres à estimer sont les suivants :

$$[(\theta_0, \theta_1, \dots, \theta_m), P_0, \sigma^2, \sigma_{\eta}^2, \mu, \sigma_1^2, v, \sigma_2^2, \rho, \sigma_3^2, \sigma_{1*}, \sigma_{2*}, \sigma_{3*}, \sigma_{12}, \sigma_{13}, \sigma_{23}].$$

La procédure itérative d'optimisation est la suivante :

- (1) Etant donné P_0 , $i \in n$, on choisit des valeurs initiales pour P_0 et σ^2 .
- (2) On calcule les D_i correspondant en résolvant numériquement (3.22).

9. On aura remarqué que puisque u_i n'apparaît pas dans (3.26), il n'est pas nécessaire de faire d'hypothèse explicite sur la distribution des u_i .

(3) On remplace les valeurs calculées pour D_i dans (3.30) et (3.31) et on maximise la fonction ainsi concentrée de vraisemblance par l'approximation quadratique ou d'autres méthodes de calcul numérique (voir Goldfeld et Quandt 1972, ch. 1).

(4) On répète l'algorithme à partir de (1) avec les valeurs estimées pour P_0 et σ^2 jusqu'à l'obtention de l'optimum.

4) *Principales caractéristiques de la méthode des écarts*

Notre nouvelle approche à l'estimation des modèles avec seuils de réaction comporte plusieurs avantages par rapport à la méthode suggérée par Dagenais :

1) Nous avons éliminé de la fonction de vraisemblance toute estimation d'intégrales multiples en évaluant à chaque itération la valeur de n intégrales simples.

2) L'estimation de la prévision conditionnelle du seuil de réaction est contrainte à l'espérance de la somme des variables d'écarts ε_1 et ε_3 , tandis que l'estimation de la prévision conditionnelle du comportement des agents qui ont pris une décision d'action est égale à la somme de la valeur limite L et des moyennes des variables d'écart des deux régimes.

3) Il n'est même pas nécessaire d'estimer explicitement le vecteur β du modèle de régression expliquant le comportement de ceux qui ont décidé d'agir. Le simple calcul de la première composante principale associée à la plus grande racine caractéristique de X suffit.

4) La méthode des écarts met beaucoup plus en évidence les déterminants du seuil de réaction comme étant les variables clefs du modèle alors que les déterminants du modèle d'action ne sont utilisés que comme indicateurs pour identifier le régime par lequel W_i est généré.

5) L'approximation de la fonction en escalier de Goldfeld et Quandt par une loi de probabilité a une similitude étonnante avec le modèle probit. Elle a, de plus, l'avantage d'estimer une valeur limite P_0 au-delà de laquelle W_i est généré par le régime II. De plus, la variance autour de cette limite détermine le degré d'approximation de la fonction en escalier par une loi de probabilité. La comparaison de P_0 avec $\mu + \rho$ devrait nous procurer une information additionnelle sur la précision de la valeur estimée pour le seuil.

6) Avec la méthode des écarts, il est impossible de faire des hypothèses aussi restrictives que celles qui sont faites par Dagenais dans sa dernière contribution (1973) sur l'erreur caractérisant la fonction d'achat et l'erreur de spécification du seuil, car il n'y a plus que l'erreur du seuil (η_i) qui entre explicitement dans le modèle.

7) Selon les résultats obtenus par Goldfeld et Quandt (1972), l'optimisation d'une fonction de vraisemblance concentrée (i.e. après approximation d'une fonction en escalier par une loi de probabilité) ne semble pas présenter de difficultés particulières au plan des méthodes de calcul numérique alors que l'optimisation de la fonction de vraisemblance de Dagenais semble être un problème de taille même pour les meilleurs experts dans les méthodes de calcul numérique. Ceci vient du fait que notre approche des variables d'écart remplace la nécessité de spécifier des lois normales tronquées pour u_i et η_i dans l'approche de Dagenais.

Jean-Guy LORANGER,
Université de Montréal.

RÉFÉRENCES

- CRAGG, J.G., UHLER, R.S., (1969), *The Demand for Automobiles*, Discussion Paper, No. 27, Department of Economics, U.B.C.
- DAGENAIS, M.G., (1969), A Threshold Regression Model », *Econometrica*, vol. 37, n° 2, pp. 193-203.
- DAGENAIS, M.G., (1973), *Application of a Threshold Regression Model to Household Purchases of Automobiles*, Cahier n° 7304, département des Sciences Economiques, Université de Montréal.
- GOLDBERGER, A.S., (1964), *Econometric Theory*, John Wiley and Sons, N.Y.
- GOLDFELD, S.M., QUANDT, R.E., (1972), *Non Linear Methods in Econometrics*, North-Holland Pub. Co., Amsterdam.
- QUANDT, R.E., (1972), « New Approach to Estimating Switching Regressions », *Journal of the American Statistical Association*, vol. 67, pp. 306-310.
- SWAMY, P.A.V.B., MEHTA, J.S., (1973), « Estimation of Switching Regressions and Its Alternatives », notes miméographiée, Federal Reserve System et Temple University.
- TOBIN, J., (1958), « Estimation of Relationships for Limited Dependent Variables », *Econometrica*, vol. 26, pp. 24-36.