

Erreurs de mesure sur les variables économiques et financières

François-Éric Racicot

WORKING PAPER
WP.2013.06

May 2013
ISSN 0701-3086

Erreurs de mesure sur les variables économiques et financières

François-Éric Racicot*

École de gestion Telfer

Université d'Ottawa

*Membre du Centre de recherche en comptabilité et en gouvernance CGA-Canada, chercheur associé à la Chaire de recherche en information financière et organisationnelle, ESG-UQAM.

Adresse : École de gestion Telfer, Université d'Ottawa, 55 rue Laurier Est, Pavillon Desmarais, Ottawa, Ontario, K1N 6N5.

Tél.: 613-562-5800 ext. 4757

Courriel : racicot@telfer.uottawa.ca

Erreurs de mesure sur les variables économiques et financières

Résumé

Cet article présente deux nouveaux estimateurs robustes en présence d'erreurs sur les variables. Ces estimateurs ont la qualité d'être beaucoup moins erratiques que leurs homologues classiques : les estimateurs de Durbin et de Pal. Ces nouveaux estimateurs sont basés sur les moments échantillonnaires d'ordre supérieur à deux. Ces estimateurs peuvent être considérés comme des types spéciaux d'estimateurs à variables instrumentales, où les instruments contiennent les variables explicatives élevées à des puissances particulières. La performance de ces nouveaux estimateurs est évaluée à partir d'une enquête canadienne sur les finances des consommateurs. Des simulations de Monte Carlo sont également effectuées et elles démontrent que les biais présentés par nos nouveaux estimateurs sont moindres que ceux associés à l'estimateur des moindres carrés ordinaires. Les estimateurs développés dans cet article participent de la nature des estimateurs de la méthode des moments généralisés (GMM) dont la popularité ne cesse de grandir. L'économétrie financière pourrait grandement bénéficier de nos nouveaux estimateurs en les appliquant à des modèles fort connus, tel le CAPM où le portefeuille du marché, qui est à la base de la vérification empirique de ce modèle, est empreint d'erreurs de mesure importantes. Les mesures de la prime de risque qui lui sont associées pourraient donc être corrigées à l'aide de nos estimateurs, cela dans le cadre de la méthode des GMM. Enfin, nous présentons une application financière où nous appliquons un des estimateurs afin d'estimer le bêta de marché, celui-ci entrant dans la formule du coût moyen du capital (WACC) que nous avons calculé et comparé à la méthode classique des MCO. Dans cette nouvelle application, nous trouvons que l'estimation du paramètre augmente significativement, ce qui est en accord avec certains faits empiriques.

Mots-clés : Estimateurs à moments supérieurs; moindres carrés ordinaires (MCO); GMM; instruments robustes; tests d'erreurs de mesure; convergence en probabilité; REQM; modèle de marché; WACC.

1. Introduction¹

La plupart des données utilisées dans les études économiques et financières empiriques contiennent des erreurs de mesure. Ces erreurs sont relativement plus importantes dans les données macro-économiques (Morgenstern, 1972; Langanskens et Van Rieckeghem, 1974), mais elles sont aussi présentes dans la plupart des études micro-économiques. Comme l'a souligné Morgenstern (1972), ce phénomène n'a jamais été pris sérieusement en considération dans le développement des techniques économétriques. Bien que dans les applications, les auteurs avertissent généralement les lecteurs que la présence possible d'erreurs de mesure peut biaiser les résultats et même conduire à la non-convergence de l'estimateur des moindres carrés ordinaires (MCO), peu d'effort particulier n'est entrepris pour atténuer ces biais et pratiquement aucun test sur la présence de ces erreurs, en l'occurrence le test bien connu d'Hausman (1978), n'est effectué. L'attitude de la plupart des chercheurs empiriques est probablement due dans bon nombre de cas au fait qu'il n'est pas toujours facile de vérifier si les variables instrumentales disponibles utilisées pour obtenir des estimateurs sans biais satisfont les conditions requises pour justifier leur utilisation (c.-à-d. à cause du problème des instruments faibles)². Dans d'autres cas, les instruments ne sont tout simplement pas accessibles aux chercheurs. Toutefois, il est important de noter que certaines recherches récentes ont été entreprises afin d'obtenir des instruments robustes (Racicot, 1993, 2000, 2003) pouvant être utilisés comme base instrumentale à l'estimateur des GMM (Racicot, 2000; Racicot et Théoret, 2008; Racicot et Théoret, 2012).

Cet article vise à donner un nouvel éclairage au phénomène des erreurs de mesure. Il se subdivise en quatre grandes parties. D'abord, de nouveaux estimateurs pour traiter les erreurs de mesure sont présentés. Ensuite, un test sur la performance de ces estimateurs est exposé. Sont également effectuées des simulations de Monte Carlo mesurant la performance relative de ces nouveaux estimateurs en regard de l'estimateur des MCO. Finalement, nous présentons une application financière du problème des erreurs de mesure. Nous généralisons

¹La réédition de l'ouvrage de Greene (2012) intègre dans plusieurs chapitres la question des erreurs de mesure et de l'endogénéité et explique l'importance que la littérature récente accorde au problème des erreurs de spécification. Ce champ de recherche semble avoir attiré l'attention de plusieurs chercheurs de l'économétrie, de la finance (Coen et Racicot, 2007; Meng, Hu et Bai, 2011; Doan et Lin, 2012) et plus particulièrement les chercheurs du domaine de l'analyse des rendements des fonds de couverture (Gregoriou, Hübner et al. (2005); Racicot et Théoret, 2008a, 2012; Lambert, 2012).

²Hahn et Hausman (2003) montrent que l'estimateur des doubles moindres carrés (2MC) est biaisé vers les MCO si les instruments sont faibles.

l'application et la définition de celles-ci dans le cadre des séries temporelles et montrons comment estimer les paramètres de modèles d'équations standards de différence en économétrie financière.

2. Conséquences des erreurs sur les variables

Les causes des erreurs sur les variables sont trop connues pour que l'on s'y attarde démesurément. Mentionnons cependant qu'il existe trois sources possibles de statistiques erronées : i) l'observateur, en faisant un choix du phénomène à observer et de sa mesure, introduit ipso facto un biais qu'il est impossible d'éviter, car un phénomène complexe ne peut jamais être décrit de façon exhaustive. Ce biais est commun à toutes les sciences et ne retient pas notre attention dans cet article; ii) l'observateur peut dissimuler de l'information pour satisfaire à ses hypothèses ou à des fins politiques; iii) les répondants peuvent mentir à l'enquêteur ou ne pas savoir la réponse exacte. Ce sera le cas, par exemple, de la richesse personnelle, où l'individu ne connaît pas exactement la valeur marchande de ses biens matériels, mais ne dispose que d'une approximation de cette valeur. Par ailleurs, il est fréquent que les variables que l'on utilise pour expliquer un phénomène ne soient pas observables ou bien qu'elles soient observées de façon erronée. À titre d'exemple, dans le modèle de Cagan (1956) sur les attentes adaptatives, les quantités offertes peuvent dépendre de variables non observables tel le prix anticipé, les anticipations étant révisées en fonction de l'erreur associée au niveau antérieur des anticipations.

Il convient maintenant de s'attarder aux conséquences des erreurs sur les variables. Le terme d'erreur est souvent représenté par une constante à laquelle s'ajoute un terme aléatoire. L'espérance mathématique de ce terme d'erreur n'est donc pas nulle. Si on utilise un modèle de régression, le terme constant de l'erreur incorporé dans la variable explicative causera un biais systématique qui n'affectera que l'estimation de la constante de la régression. Comme cette estimation est souvent de moindre intérêt, on ne s'intéressera ici qu'à la partie aléatoire de l'erreur, qu'on supposera de moyenne nulle. L'exemple suivant illustre ce problème. Soit :

$$y = \alpha + \tilde{X}\beta + e \quad (1)$$

où $X = \tilde{X} + w$ et $w = E(w) + w - E(w)$, $E(w) = c$, $v = w - E(w)$ et $w = c + v$, où c est la partie constante de l'erreur (w) et v la partie aléatoire. En faisant les substitutions appropriées, on obtient :

$$\begin{aligned} y &= \alpha + X\beta + e - w\beta \\ &= \alpha + X\beta + e - c\beta - v\beta \\ &= \alpha - c\beta + X\beta + e - v\beta \end{aligned} \quad (2)$$

Certes, $c\beta$ est la partie qui biaise cette estimation (ce que l'on obtient est une estimation de $\alpha - c\beta$). Mais beaucoup plus intéressante est l'estimation du paramètre β . On suppose que $X = \tilde{X} + v$ et $E(v) = 0$.³

Un premier résultat ayant trait aux conséquences des erreurs de mesure est le suivant. S'il y a présence d'erreurs de mesure aléatoires sur les variables dépendantes, ces erreurs se combineront à l'erreur de la régression sans créer de biais d'estimation. Mais dans le cas où les variables explicatives sont affectées d'erreurs de mesure, les conséquences sont alors tout autres.

Un modèle de régression linéaire univarié sans terme constant illustre bien ce problème.⁴ Soit :

$$y = \tilde{X}\beta + e \quad (3)$$

Dans cette équation, y est un vecteur de variables dépendantes de dimension $N \times 1$ et \tilde{X} , un vecteur de variables explicatives de dimension $N \times I$ avec $E(\tilde{X}_i^2) = \sigma_{\tilde{X}}^2$. e est un vecteur d'erreurs de dimension $N \times 1$ avec $E(e) = 0$ et $E(ee') = \sigma_e^2 I$, $E(e_i^2) = \sigma_e^2$. Enfin β est un paramètre à estimer. On suppose que les vraies variables \tilde{X} ne sont pas observables. On observe plutôt :

$$X = \tilde{X} + v \quad (4)$$

³Les problèmes de corrélation entre les erreurs d'observation ne sont pas considérés. On se contentera de traiter le problème plus simple des erreurs de mesure qui ne sont pas autocorrélées. Toutefois, il ne faut pas confondre ce type d'autocorrélation avec celui de l'autocorrélation des résidus de la régression. Dans ce cas particulier, Dagenais (1994) montre qu'en utilisant une méthode classique de correction de l'autocorrélation (e.g. Cochrane et Orcutt, 1949; Prais et Winsten, 1954) en présence d'erreurs de mesure le problème pourrait être amplifié.

⁴ Pour le cas multivarié, voir l'appendice A.

où v est un vecteur d'erreurs de mesure avec $E(v) = 0$ et $E(v_i^2) = \sigma_v^2$. On suppose aussi que : $E(v e') = E(\tilde{X} v') = E(\tilde{X} e') = 0$. En remplaçant \tilde{X} par $X - v$ dans l'équation de base, on obtient :

$$\begin{aligned} y &= X\beta - \beta v + e \\ &= X\beta + e^* \end{aligned} \quad (5)$$

où $e^* = e - \beta v$. Si on applique les M.C.O. sur la dernière équation, on a :

$$\hat{\beta} = \beta + (X'X)^{-1} X' e^*$$

L'estimateur qui en résulte est non convergent puisque :

$$plim(\hat{\beta}) = \beta + plim(X'X/N)^{-1} plim(X' e^* / N) \quad (6)$$

En utilisant les théorèmes sur le calcul de la limite en probabilité ($plim$) apparaissant à l'appendice B, on a, sous les hypothèses appropriées :

$$plim(X'X/N) = E_{N \rightarrow \infty}(X'X/N)$$

et

$$plim(X' e^* / N) = E_{N \rightarrow \infty}(X' e^* / N).$$

On développe ces deux expressions comme suit :

$$\begin{aligned} E_{N \rightarrow \infty}(X'X/N) &= E_{N \rightarrow \infty} [\sum_{i=1}^N (\tilde{X}_i + v_i)(\tilde{X}_i + v_i)] / N = E_{N \rightarrow \infty} [\sum_{i=1}^N (\tilde{X}_i^2 + 2\tilde{X}_i v_i + v_i^2)] / N \\ &= \sigma_{\tilde{X}}^2 + \sigma_v^2 \end{aligned}$$

$$E_{N \rightarrow \infty}(X' e^* / N) = E_{N \rightarrow \infty} \left[\sum_{i=1}^N (\tilde{X}_i + v_i)(e_i - \beta v_i) / N \right] = -\beta \sigma_v^2$$

et, par conséquent,

$$plim(\hat{\beta}) = \beta - \beta \frac{\sigma_v^2}{\sigma_{\tilde{X}}^2 + \sigma_v^2} = \frac{\beta}{1 + \lambda} \quad (7)$$

où :

$$\lambda = \frac{\sigma_v^2}{\sigma_{\tilde{X}}^2} \quad (8)$$

On peut conclure que les MCO sont asymptotiquement biaisés et non convergents et que le biais dépend de λ , soit de l'importance relative de la variance des erreurs de mesure sur les variables explicatives. En effet, λ est défini comme le ratio de la variance des erreurs de mesure à la variance des variables explicatives non observées (\tilde{X}). λ peut être vu aussi comme l'importance relative des erreurs de mesure dans les variables explicatives.

Le second résultat ayant trait aux conséquences des erreurs de mesure est associé à la fonction de vraisemblance. Quel est en effet l'impact des erreurs de mesure dans les variables explicatives sur les estimateurs du maximum de vraisemblance ? Pour répondre à cette interrogation, reconsidérons le modèle de régression linéaire univariée présenté plus haut et supposons que e et v sont normalement et indépendamment distribués (i.i.d.), soit :

$$E(e) = E(v) = 0, E(ev') = 0, E(ee') = \sigma_e^2 I, E(e_i^2) = \sigma_e^2, E(vv') = \sigma_v^2 I, E(v_i^2) = \sigma_v^2.$$

Le paramètre β ne peut pas être estimé de façon convergente par le maximum de vraisemblance car il y a carence d'information pour identifier le paramètre. Le problème de pénurie d'information peut être détecté lorsqu'on applique la méthode du maximum de vraisemblance (Judge et al., 1985). La fonction de densité conjointe de e et v est :

$$f(e, v) = (2\pi \sigma_e^2)^{-N/2} \exp[-(e'e)/2 \sigma_e^2] \times (2\pi \sigma_v^2)^{-N/2} \exp[-(v'v)/2 \sigma_v^2] \quad (9)$$

et à l'aide du Jacobien on trouve la fonction de vraisemblance de Y et X soit :

$$L(y, X; \beta, \sigma_e^2, \sigma_v^2, \tilde{X}) = (2\pi \sigma_e^2)^{-N/2} (2\pi \sigma_v^2)^{-N/2} \exp[-((y - X\beta)'(y - X\beta))/2 \sigma_e^2] \exp[-((X - \tilde{X})'(X - \tilde{X}))/2 \sigma_v^2]. \quad (10)$$

Le logarithme de cette fonction est égal à :

$$\begin{aligned} \ln L(y, X; \beta, \sigma_e^2, \sigma_v^2, \tilde{X}) = & -N(\ln 2\pi) - N/2(\ln \sigma_e^2 + \ln \sigma_v^2) - (X - \tilde{X})'(X - \tilde{X})/2 \sigma_v^2 - \dots \\ & \dots - (y - \tilde{X}\beta)'(y - \tilde{X}\beta)/2 \sigma_e^2 \end{aligned} \quad (11)$$

Si l'on dérive L par rapport aux paramètres recherchés, on obtient :

$$\begin{aligned} \partial L / \partial \beta &= 0 \Rightarrow \tilde{X}^{*'}(y - \tilde{X}^* \beta^*) = 0 \\ \partial L / \partial \sigma_v^2 &= 0 \Rightarrow (X - \tilde{X}^*)'(X - \tilde{X}^*) = N \sigma_v^2 \\ \partial L / \partial \sigma_e^2 &= 0 \Rightarrow (y - \tilde{X}^* \beta^*)'(y - \tilde{X}^*) = N \sigma_e^{2*} \\ \partial L / \partial \tilde{x} &= 0 \Rightarrow (X - \tilde{X}^*) \sigma_e^{2*} + \beta^* (y - \tilde{X}^* \beta^*) \sigma_v^{2*} = 0 \end{aligned}$$

et si l'on résout le système d'équations, on a :

$$\sigma_v^{2*} = \beta^{2*} \sigma_e^{2*}, \quad (12)$$

ce qui n'est pas un résultat satisfaisant (voir l'appendice A). L'incapacité de la méthode du maximum de vraisemblance de fournir une seule estimation est le résultat d'un manque d'information. Le système est sous-identifié.

3. Les estimateurs

3.1. Les estimateurs usuels

Pour obtenir des estimateurs convergents en présence d'erreurs de mesure au chapitre des variables explicatives, une méthode préconisée dans la plupart des économètres est celle des variables instrumentales. Cette méthode usuelle pour générer des estimateurs convergents en présence d'erreurs de mesure consiste à trouver une matrice d'instruments pour éliminer le problème de corrélation entre la matrice des variables explicatives et les erreurs de régression. Il est cependant très difficile de trouver une telle matrice.

Si l'on reprend le modèle de régression linéaire présenté au chapitre précédent avec les mêmes hypothèses, soit :

$$y = X\beta + e \quad (13)$$

et qu'on suppose en premier lieu qu'il n'y a pas d'erreur sur les variables explicatives, alors en appliquant

$$\hat{\beta} = (X'X)^{-1}(X'y) \quad (14)$$

les MCO sur cette équation, on trouve :

et si on substitue y par sa valeur :

$$\hat{\beta} = \beta + (X'X)^{-1}X'e. \quad (15)$$

Pour montrer que $\hat{\beta}$ est convergent, on calcule le $p \lim(\hat{\beta})$:

$$p \lim \hat{\beta} = \beta + (p \lim(X'X/N))^{-1} p \lim(X'e/N) = \beta + Q^{-1} \times 0 \quad (16)$$

où $p \lim(X'X/N) = \lim_{N \rightarrow \infty} (X'X/N) = Q$, une matrice non singulière, une hypothèse bien connue du modèle classique linéaire général, et $p \lim(X'e/N) = \lim_{N \rightarrow \infty} E(X'e/N)$ où $E(X'e) = 0$ car X n'est pas stochastique et $E(e) = 0$, une autre hypothèse du modèle classique. Le fait que $E(X'e) = 0$ nous indique que les vecteurs de X et de e ne sont pas orthogonaux. Mais dans le cas où $p \lim(X'e/N) \neq 0$ (orthogonalité entre X et e), les MCO deviennent automatiquement non convergents. La méthode des variables instrumentales consiste donc à trouver une matrice, par exemple W , telle que $p \lim(W'e/N) = 0$.

L'estimateur des MCO devient, si on remplace la matrice X par W :

$$\hat{\beta} = (W'X)^{-1}W'y \quad (17)$$

où $p \lim \hat{\beta} = \beta$. La matrice W est formée de variables fortement corrélées avec celles de X mais non pas avec e . En pratique, il est très difficile de trouver une matrice, comme la matrice W , qui possède toutes les caractéristiques voulues.

Une alternative à l'approche des variables instrumentales mais qui n'a reçu que très peu d'attention dans la littérature est discutée en particulier par Pal (1980). Il présente, pour un modèle de régression qui ne contient qu'une seule variable explicative, six estimateurs basés sur les moments du troisième ordre des observations. Ces estimateurs ont, sous des hypothèses très raisonnables, la propriété d'être convergents même s'il y a présence d'erreurs de mesure sur les variables explicatives. Il mentionne que l'un de ces estimateurs a été suggéré par Durbin (1954), un autre par Drion (1951) et un troisième par Geary (1942). L'un des estimateurs qui retient notre attention est celui de Durbin (β_d). Cet estimateur a la propriété d'être aussi sans biais lorsqu'il n'y a pas d'erreurs sur les variables explicatives, ce qui n'est pas le cas pour les autres estimateurs mentionnés par Pal. Il fait allusion au fait que l'estimateur de Durbin peut être facilement généralisé au cas où le modèle contiendrait plus d'un régresseur.

Il existe aussi d'autres estimateurs convergents, même en présence d'erreurs de mesure sur les variables explicatives, basés sur des moments échantillonnaires d'ordre plus élevé. Pour le cas où un modèle de régression aurait seulement une variable explicative, des estimateurs basés sur les quatrièmes moments ont été proposés par Geary (1942) et Pal (1980). L'un de ces estimateurs sera également considéré, car il demeure sans biais dans le cas où il n'y a pas

d'erreurs du côté des variables indépendantes. Cet estimateur, de même que β_d , sera généralisé au cas où il y a plus d'une variable explicative identifiée par β_p .

β_d et β_p dans leur forme univariée (forme initiale) sont tels que :

$$\beta_d = \frac{\sum x_i^2 y_i}{\sum x_i^3} \quad (18)$$

$$\beta_p = \frac{\sum x_i^3 y_i - 3(\sum x_i^2 / N) \sum x_i y_i}{\sum x_i^4 - 3(\sum x_i^2 / N) \sum x_i^2}$$

On peut facilement démontrer que si les erreurs de mesure sont indépendantes et suivent une loi normale, $p \lim \beta_d = \beta$, $p \lim \beta_p = \beta$. Le modèle de régression univarié sous-jacent s'écrit :

$$y = \alpha + \tilde{X}\beta + u \quad (19)$$

$X = \tilde{X} + v$, $v \sim N(0, \sigma_v^2 I)$, $E(u) = 0$, $V(u) = \sigma_u^2 I$, $V(u_i) = \sigma_u^2$ et les variables x et y , telles qu'exprimées plus haut, sont en déviation de la moyenne. Les particularités de β_p sont les termes de correction $3(\sum x_i^2 / N) \sum x_i y_i$ du numérateur et de $3(\sum x_i^2 / N) \sum x_i^2$ du dénominateur, introduits pour tenir compte du fait que les erreurs v sont normalement distribuées.

Sous l'hypothèse H0 (absence d'erreur de mesure sur les variables explicatives), β_d et β_p sont sans biais. Mais ces estimateurs sont beaucoup plus erratiques que les estimateurs des MCO correspondants (voir : Kendall et Stuart, 1963; Malinvaud, 1978). Ceci explique probablement le fait que ces estimateurs ne soient plus utilisés dans les applications actuelles.

De nouveaux estimateurs sont maintenant répertoriés (Dagenais et Dagenais, 1991; Racicot, 1993) basés sur les moments échantillonnaires d'ordre supérieur à deux, qu'on appellera β_H et β_E , β_E étant une extension de β_H . β_H est une combinaison linéaire matricielle pondérée de la version généralisée de β_d et β_p (Dagenais et Dagenais, 1991⁵; Racicot, 1993⁶). Cet estimateur, en plus d'être sans biais sous H0 et convergent sous l'alternative H1 (voir l'appendice B pour une démonstration), a la qualité d'être beaucoup moins erratique que β_d ou

⁵ Voir Dagenais et Dagenais (1997).

⁶ Voir Racicot (2000, 2003).

β_p pris séparément Comme il a été mentionné par Pal (1980), tous ces estimateurs peuvent être considérés comme des types spéciaux d'estimateurs à variables instrumentales, où les instruments contiennent les variables explicatives élevées à des puissances particulières. Ces estimateurs participent de la nature des estimateurs de la méthode des moments généralisés (GMM) dont la popularité ne cesse de grandir depuis que Hansen les a introduits en 1982 (Hansen, 1982).

Le test de l'hypothèse H_0 peut être fait avec l'estimateur β_H . Mais il est à noter que le même test peut être effectué de deux autres façons. L'une d'elles est analogue au test proposé par Hausman (1978). L'autre se fonde sur la propriété d'indépendance entre les erreurs de la régression et les régresseurs, sous H_0 . Le test en question est exact dans les petits échantillons car sous H_0 l'estimateur β_H est sans biais (la distribution exacte de la statistique est dérivable). Il apparaît d'après les expériences effectuées, que le test proposé peut être utile pour détecter la présence d'erreurs dans les variables explicatives dans les grands échantillons, même si ces erreurs sont petites et que le coefficient de corrélation multiple est faible. Le test peut aussi faire montre de puissance dans les petits échantillons quand le coefficient de corrélation multiple est élevé et les erreurs de mesure, importantes. Dans le cas où l'on suspecte fortement la présence d'erreurs de mesure sur les variables explicatives, il est possible d'utiliser la version étendue de β_H , soit β_E , qui est aussi convergente et asymptotiquement plus efficace quand H_1 est acceptée.

3.2. L'estimateur β_H

Soit le modèle de régression suivant :

$$y = \alpha i + \tilde{X}\beta + u \quad (20)$$

où \tilde{X} est une matrice de dimension $N \times K$ de variables non aléatoires. Cette matrice est celle des variables explicatives mesurées sans erreur et on suppose que $\lim_{N \rightarrow \infty} \frac{\tilde{X}\tilde{X}'}{N} = Q$ où Q est finie et non singulière. Le vecteur u est un vecteur d'erreurs résiduelles normalement distribuées, de dimension $N \times 1$ avec $E(u) = 0$, $E(uu') = \sigma_u^2 I_N$. y est un vecteur d'observations sur la variable dépendante de dimension $N \times 1$. Enfin β est un vecteur de paramètres à estimer de dimension $K \times 1$, α est la constante à estimer et i est un vecteur unité

de dimension $N \times I$. $\text{var}(u) = \sigma_u^2 I_N$ implique que les erreurs de la régression ne sont pas corrélées entre elles et que la variance est homoscédastique.

On suppose que \tilde{X} est non observable et que la matrice X est observée à la place, soit :

$$X = \tilde{X} + v \quad (21)$$

où v est une matrice d'erreurs de mesure (normalement distribuées) sur les variables explicatives, de dimension $N \times K$.

On suppose aussi que $\text{var}[\text{vech}(v)] = \Sigma \otimes I_N$ où $\text{var}[\cdot]$ signifie: matrice de covariances et Σ est une matrice symétrique, semi-définie positive. Soit :

$$\Sigma = \begin{pmatrix} \sigma_{11} & \cdot & \cdot & \sigma_{1K} \\ \sigma_{21} & \sigma_{22} & \cdot & \sigma_{2K} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_{K1} & \cdot & \cdot & \sigma_{KK} \end{pmatrix} \quad \text{et} \quad \text{var}[\text{vech}(v)] = \begin{pmatrix} \sigma_{11} I_N & \cdot & \cdot & \sigma_{1K} I_N \\ \sigma_{21} I_N & \sigma_{22} I_N & \cdot & \sigma_{2K} I_N \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_{K1} I_N & \cdot & \cdot & \sigma_{KK} I_N \end{pmatrix} \quad (22)$$

Comme on peut le constater ci-dessus, les hypothèses émises sur les erreurs de mesure des variables explicatives impliquent que les erreurs sont indépendantes entre observations, soit : $\text{cov}(v_{ij}, v_{i+k, j'}) = 0, i = 1, \dots, N, j, j' = 1, \dots, K, k \neq 0$, mais qu'elles sont dépendantes entre variables pour une même observation, soit: $\text{cov}(v_i, v_j) \neq 0, i \neq j$. Par exemple: la matrice de covariance des erreurs de mesure de la variable X_i est : $\sigma_{11} I_N$. Cette hypothèse implique aussi que pour une même variable, les variances des erreurs de mesure sont homoscédastiques.

β_p n'a pas été généralisé pour le cas où il y aurait plus d'une variable explicative (la généralisation de β_d est donnée dans Pal (1980)). La version multivariée de ces estimateurs (Dagenais et Dagenais, 1992b; Racicot, 1993) qui servent à la construction de β_H et β_E , sous forme d'estimateurs à variables instrumentales, se présente comme suit :

$$\beta_D = (z_1' x)^{-1} z_1' y \quad \text{où} \quad z_1 = \begin{pmatrix} x_{11}^2 & \cdot & \cdot & x_{1K}^2 \\ x_{21}^2 & x_{22}^2 & \cdot & x_{2K}^2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{N1}^2 & \cdot & \cdot & x_{NK}^2 \end{pmatrix}. \quad (23)$$

Les x_{ij} sont les éléments de la matrice x et $x = AX$ où $A = I_N - ii'/N$. La matrice x correspond à la matrice X calculée en déviation de la moyenne. On applique la même procédure pour y où $y = AY$.

$$\beta_p = (z_2' x)^{-1} z_2' y \text{ où } z_2' = z_3' - 3D(x'x/N)x' \text{ et } z_3 = \begin{pmatrix} x_{11}^3 & \cdot & \cdot & x_{1K}^3 \\ x_{21}^3 & x_{22}^3 & \cdot & x_{2K}^3 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{N1}^3 & \cdot & \cdot & x_{NK}^3 \end{pmatrix} \quad (24)$$

Encore une fois les éléments de x et y sont exprimés en déviation de la moyenne, $D(x'x/N)$ est une matrice diagonale de dimension $K \times K$ avec sur la diagonale les éléments correspondants de la matrice $x'x/N$. Pour calculer $D(x'x/N)$, on multiplie chaque élément de $x'x/N$ par l'élément correspondant de la matrice identité. Soit: $x = (x_1 x_2 \dots x_k)$, où les éléments de x sont des vecteurs colonnes.

Donc,

$$D(x'x/N) = \begin{pmatrix} x_1'x_1/N & \cdot & \cdot & 0 \\ 0 & x_2'x_2/N & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & x_k'x_k/N \end{pmatrix} \quad (25)$$

Connaissant maintenant β_D et β_p , on peut dériver β_H :

$$\beta_H = W \begin{pmatrix} \beta_D \\ \beta_p \end{pmatrix} \text{ où } W = (C' S^{-1} C)^{-1} C' S^{-1}, C = \begin{pmatrix} I_K \\ I_K \end{pmatrix} \quad (26)$$

et S est la matrice de covariances de $\begin{pmatrix} \beta_D \\ \beta_p \end{pmatrix}$ sous H_0 . Plus précisément, β_H est obtenu en appliquant les MCG sur une combinaison de β_D et β_p ; β_D est sans biais sous H_0 et a comme matrice de covariance S_D , soit :

$$\beta_D = (z_1' x)^{-1} z_1' y \quad (27)$$

et,

$$y = x\beta + Au \text{ où } y = AY \text{ et } x = AX. \quad (28)$$

Si l'on substitue dans β_D y par sa valeur, on obtient :

$$\beta_D = (z_1'x)^{-1} z_1'x\beta + (z_1'x)^{-1} z_1'Au \quad (29)$$

alors pour montrer que β_D est sans biais sous H_0 , on calcule :

$$E(\beta_D) = \beta + (z_1'x)^{-1} z_1'AE(u) \text{ et } E(u) = 0 \text{ donc } E(\beta_D) = \beta.$$

La matrice de covariances de β_D est :

$$\begin{aligned} S_D &= E[(\beta_D - \beta)(\beta_D - \beta)'] \\ &= E[((\beta + (z_1'x)^{-1} z_1'Au) - \beta)((\beta + (z_1'x)^{-1} z_1'Au) - \beta)'] \\ &= E[((z_1'x)^{-1} z_1'Au)((z_1'x)^{-1} z_1'Au)'] \\ &= E[(z_1'x)^{-1} z_1'Auu'A' z_1(x' z_1)^{-1}] \\ &= (z_1'x)^{-1} z_1'A z_1(x' z_1)^{-1} \sigma_u^2 \text{ car } AA' = A, E(uu') = \sigma_u^2 I_N. \end{aligned}$$

On procède de la même façon pour β_P :

$$\beta_P = (z_2'x)^{-1} z_2'y, \quad y = AY, \quad x = AX \quad (30)$$

En substituant y dans β_P , on trouve : $\beta_P = \beta + (z_2'x)^{-1} z_2'Au$ et donc β_P est sans biais sous H_0 car $E(\beta_P) = \beta$. La matrice de covariance de β_P est :

$$S_P = E[(\beta_P - \beta)(\beta_P - \beta)'] = (z_2'x)^{-1} z_2'A z_2(x' z_2)^{-1} \sigma_u^2.$$

Maintenant, pour obtenir β_H dérivé de β_P et β_D , il faut appliquer les MCG sur la forme suivante :

$$\begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} = C\beta + \begin{pmatrix} u_D \\ u_P \end{pmatrix} \text{ où } C = \begin{pmatrix} I_K \\ I_K \end{pmatrix}, \quad u_D = (z_1'x)^{-1} z_1'Au \text{ et } u_P = (z_2'x)^{-1} z_2'Au$$

ce qui donne :

$$\beta_H = (C' S^{-1} C)^{-1} C' S^{-1} \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} \quad (31)$$

ou plus précisément, si on applique les M.C.O sous la forme suivante:

$$P \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} = PC\beta + P \begin{pmatrix} u_D \\ u_P \end{pmatrix} \quad (32)$$

où P est une matrice non singulière de dimension $2K \times 2K$ telle que $P'P = S^{-1}$, alors :

$$(PC)'P \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} = (PC)'PC \beta_H$$

$$\beta_H = (C'P'PC)^{-1} (C'P'P) \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} \quad (33)$$

et donc

$$\beta_H = (C' S^{-1} C)^{-1} C' S^{-1} \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} \quad (34)$$

où S est la matrice de covariance de $\begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix}$ et est donnée par: $\begin{pmatrix} S_D & S_{DP} \\ S_{DP}' & S_P \end{pmatrix}$

où :

$$S_{DP} = (z_1' x)^{-1} z_1' A z_2 (x' z_2)^{-1} \sigma_u^2. \quad (35)$$

La matrice de covariance de β_H est donnée par :

$$V(\beta_H) = (C' S^{-1} C)^{-1} (\sigma_u^2 \text{ est inclus dans } S^{-1})$$

β_H sera sans biais car β_D et β_P sont sans biais sous H_0 soit :

$$E(\beta_H) = (C' S^{-1} C)^{-1} C' S^{-1} E \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix}$$

$$= (C' S^{-1} C)^{-1} C' S^{-1} C\beta \text{ car } E \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} = c\beta$$

$$= \beta \quad (36)$$

L'application des MCG sous la forme décrite plus haut fournit un estimateur qui est une combinaison linéaire matricielle optimale de β_D et β_P . Il est possible de montrer que l'estimateur résultant d'un arrangement de la sorte aura une variance plus petite ou égale à la plus petite des variances des deux estimateurs β_D ou β_P (théorème de Theil, H. et A.S. Goldberger, 1961). On peut montrer que β_H demeure convergent (convergence en probabilité) en présence d'erreurs de mesure sur les variables explicatives (voir appendice B).

Il existe une autre façon (plus simple) pour déduire β_H , ceci en utilisant la méthode des régressions artificielles (MacKinnon, 1992). Il s'agit, dans un premier temps, de former K régressions artificielles en prenant X comme variable dépendante et (z_1, z_2) comme variables indépendantes soit :

$$X = (i, z_1, z_2)\Gamma + h \quad (37)$$

où

$$\Gamma = \begin{pmatrix} \gamma_A' \\ \gamma_B' \\ \gamma_C' \end{pmatrix}, \text{ explicitement : } \begin{pmatrix} \gamma_{A_{11}} & \gamma_{A_{12}} & \dots & \gamma_{A_{1K}} \\ \gamma_{B_{11}} & \gamma_{B_{12}} & \dots & \gamma_{B_{1K}} \\ \cdot & \cdot & \dots & \cdot \\ \gamma_{B_{K1}} & \gamma_{B_{K2}} & \dots & \gamma_{B_{KK}} \\ \gamma_{C_{11}} & \gamma_{C_{12}} & \dots & \gamma_{C_{1K}} \\ \cdot & \cdot & \dots & \cdot \\ \gamma_{C_{K1}} & \gamma_{C_{K2}} & \dots & \gamma_{C_{KK}} \end{pmatrix} \quad (38)$$

et :

$$h = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1K} \\ h_{21} & h_{22} & \dots & h_{2K} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ h_{N1} & h_{N2} & \dots & h_{NK} \end{pmatrix} \quad (39)$$

est une matrice d'erreurs de dimension $N \times K$ avec $E(h) = 0$. On applique les MCO sur $X = (i, z_1, z_2)\Gamma + h$ et on trouve ainsi $\hat{\Gamma}$ et on calcule $\hat{X} = (i, z_1, z_2)\hat{\Gamma}$. Pour terminer, on transforme \hat{X} en \hat{x} (en déviation de la moyenne) et on utilise \hat{x} comme matrice d'instruments, soit :

$$\beta_H = (\hat{x}'x)^{-1} \hat{x}'y. \quad (40)$$

On peut remplacer x par \hat{x} , soit :

$$\beta_H = (\hat{x}'\hat{x})^{-1} \hat{x}'y; \quad (41)$$

Ces transformations reposent sur les égalités suivantes :

$$x = (z_1, z_2)' \hat{\Gamma} + \hat{h} \quad \text{et} \quad \hat{x}'\hat{h} = 0. \quad (42)$$

Enfin, la matrice de covariance de β_H sous H_0 est :

$$V(\beta_H) = (\hat{x}'\hat{x})^{-1} \sigma_u^2 \quad \text{où} \quad \sigma_u^2 \quad \text{est estimé par} \quad \hat{\sigma}_u^2 = (u_H' u_H / N - K - 1) \quad \text{et} \quad u_H = y - x\beta_H$$

Connaissant maintenant β_H on peut passer à la présentation de β_E comme il a été mentionné plus haut dans cet article.

3.3. L'estimateur β_E

L'estimateur β_H , dérivé d'une combinaison linéaire matricielle optimale de β_D et β_P sous H_0 , demeure convergent lorsqu'il y a des erreurs sur les variables explicatives. Mais il est possible de déduire un estimateur alternatif (β_E) qui serait une combinaison linéaire matricielle asymptotiquement optimale de β_D et β_P sous l'hypothèse H_1 (présence d'erreurs de mesure sur les variables explicatives). On peut trouver cet estimateur en substituant dans β_H la matrice \hat{S}^* à la matrice S . Soit :

$$\beta_E = W^* \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} \quad (43)$$

$$\text{où} \quad W^* = (C' \hat{S}^{*-1} C)^{-1} C' \hat{S}^{*-1} \quad \text{et} \quad \hat{S}^* = \begin{pmatrix} \hat{S}_D^* & \hat{S}_{DP}^* \\ \hat{S}_{DP}^{*'} & \hat{S}_P^* \end{pmatrix}.$$

Le calcul de la matrice \hat{S}^* s'effectue comme suit sous H_1 :

$$\Psi = \sqrt{N} (\beta_D - \beta) = (z_1' x / N)^{-1} z_1' A \eta / \sqrt{N}. \quad (44)$$

Ψ a la même distribution asymptotique que $\Lambda = [plim(z_1'x/N)]^{-1} z_1' A \eta / \sqrt{N}$, puisque $plim(\Lambda - \Psi) = 0$. Donc la matrice asymptotique de covariance de $\sqrt{N}(\beta_D - \beta)$ est :

$$\begin{aligned} \lim_{N \rightarrow \infty} V(\Lambda) &= \lim_{N \rightarrow \infty} E(\Lambda \Lambda') = \lim_{N \rightarrow \infty} E([plim(z_1'x/N)]^{-1} z_1' A \eta / \sqrt{N})([plim(z_1'x/N)]^{-1} z_1' A \eta / \sqrt{N})' \\ &= ([plim(z_1'x/N)]^{-1} \lim_{N \rightarrow \infty} E \frac{z_1' A \eta \eta' A z_1}{N} [plim(z_1'x/N)]^{-1}). \end{aligned} \quad (45)$$

Cette expression peut être estimée de façon convergente (White, 1980) par :

$$N \hat{S}_D^* = [z_1'x/N]^{-1} \frac{\sum_{i=1}^N \tilde{z}_{i1}' \eta_{Hi}^2 \tilde{z}_{i1}}{N} [x' z_1/N]^{-1} \quad (46)$$

où $\tilde{z}_{i1} = z_{i1} - \bar{z}_1$, z_{i1} est la i ème ligne de z_1 et \bar{z}_1 est un vecteur ($1 \times 1K$) de moyennes des K variables contenues dans z_1 , $\eta_{Hi} = y_i - x_i \beta_H$ est le i ème élément de η_H , y_i et x_i sont respectivement les éléments de la i ème ligne de y et x . De la même façon on trouve que :

$$N \hat{S}_P^* = [z_2'x/N]^{-1} \frac{\sum_{i=1}^N \tilde{z}_{i2}' \eta_{Hi}^2 \tilde{z}_{i2}}{N} [x' z_2/N]^{-1} \quad (47)$$

et

$$N \hat{S}_{DP}^* = [z_1'x/N]^{-1} \frac{\sum_{i=1}^N \tilde{z}_{i1}' \eta_{Hi}^2 \tilde{z}_{i2}}{N} [x' z_2/N]^{-1}. \quad (48)$$

3.4. Présentation des tests

L'approche des régressions artificielles utilisée pour dériver β_H et β_E peut être utilisée pour tester l'hypothèse H_0 . On définit :

$$\hat{h} = X - \hat{X}^7 \quad (49)$$

où $\hat{X} = (i, z_1, z_2) \hat{\Gamma}$ et $X = (i, z_1, z_2) \hat{\Gamma} + \hat{h}$. Si l'on suppose qu'il n'y a pas d'erreur sur les variables explicatives ($X = \tilde{X}$), on peut écrire :

$$\begin{aligned} y &= \alpha i + X\beta + u = \alpha i + (\hat{h} + \hat{X})\beta + u \\ &= \alpha i + \hat{h}\beta + \hat{X}\beta + u \end{aligned} \quad (50)$$

⁷Cette matrice peut s'interpréter comme étant la matrice des variables filtrées pour les nonlinéarités. Elle est également définie comme étant la matrice des 'w': $\hat{w} = X - \hat{X}$.

si l'on postule $\beta = \gamma \Rightarrow y = \alpha i + \hat{X}\beta + \hat{h}\gamma + u$,

on peut écrire aussi que :

$$\begin{aligned} y &= \alpha i + \hat{X}\beta + \hat{h}\beta + \hat{h}\gamma - \hat{h}\beta + u \\ &= \alpha i + \hat{X}\beta + \hat{h}\beta + \hat{h}(\gamma - \beta) + u \\ &= \alpha i + \hat{X}\beta + \hat{h}\beta + \hat{h}(\theta) + u \end{aligned}$$

où $\theta = \gamma - \beta = 0$:

$$\begin{aligned} &= \alpha i + (\hat{X} + \hat{h})\beta + \hat{h}(\theta) + u \\ &= \alpha i + X\beta + \hat{h}\theta + u. \end{aligned} \tag{51}$$

Alors sous H0, on régresse Y sur X et \hat{h} ⁸. Le résultat attendu est que les éléments du vecteur de coefficients associés à la matrice \hat{h} soient tous égaux à zéro. Le test F habituel, qui a $(K, N - 2K - 1)$ degrés de liberté (K restrictions et $2K + 1$ paramètres dans la régression), est utilisé pour tester l'hypothèse nulle ($H_0 : \theta = 0$).

Deux façons alternatives existent pour tester exactement le même phénomène :

i) Un test qui repose sur le fait que, sous H0, β_H et β_L sont sans biais donc $E(\beta_H - \beta_L) = 0$, alors que sous H1, $plim(\beta_H - \beta_L) \neq 0$.

Le test est donné par :

$$m = (\beta_H - \beta_L)' G^{-1} (\beta_H - \beta_L) / \hat{\sigma}^2 \tag{52}$$

où $G \hat{\sigma}^2$ est l'estimateur de la matrice des covariances de $(\beta_H - \beta_L)$ sous H0. On calcule cette matrice de covariances par $V[\beta_H - \beta_L] = E[(\beta_H - \beta_L)(\beta_H - \beta_L)'] = G \sigma^2$ et $G = J'J$

$$\text{où } J = (C' S^{-1} C)^{-1} C' S^{-1} \begin{pmatrix} (z_1' x)^{-1} z_1' A \\ (z_2' x)^{-1} z_2' A \end{pmatrix} - (x' x)^{-1} x'.$$

On estime σ^2 par $\hat{\sigma}^2 = \frac{e_L' e_L}{N - K - 1}$ où $e_L = y - x \beta_L$. On peut estimer aussi $\hat{\sigma}^2$ par $\frac{e_H' e_H}{N - K - 1}$

où $e_H = y - x \beta_H$.

⁸ À noter que cette façon de procéder pour tester les erreurs de mesures est celle des régressions artificielles (ou régressions auxiliaires). Voir Johnston et Dinardo (1997); Pindyck et Rubinfeld (1998); Heij et al. (2004); Gujarati et Porter (2009); Greene (2012).

Enfin, on calcule exactement la "p-value" associée à m . Cette valeur est définie comme étant la probabilité que sous H_0 la variable aléatoire correspondante à m ait une valeur supérieure à celle observée pour l'échantillon considéré. Pour un test effectué à un niveau de 5 % de signification, H_0 sera rejeté si la "p-value" de m est plus petite que 0,05 (appendice B pour le calcul de la "p-value" associée à m).

ii) La deuxième façon est, sous H_0 , de faire la régression de e_H dérivé de β_H sur x soit : $e_H = x\theta + \xi$ où ξ est un vecteur de dimension $N \times 1$ d'erreurs de régression. θ est un vecteur de dimension $K \times 1$ de paramètres inconnus. Sous H_0 , la corrélation entre e_H et x est nulle, donc tous les éléments de θ devraient être égaux à zéro. S'il y avait des erreurs dans les variables explicatives, les éléments de θ seraient significativement différents de zéro. On estime ici θ par $\hat{\theta}$ en utilisant les MCG, car sous H_0 :

$$\begin{aligned} \xi &= e_H \\ e_H &= y - x\beta_H \\ &= x\beta + Au - x\beta + x(C'S^{-1}C)^{-1}C'S^{-1} \begin{pmatrix} (z_1'x)^{-1}z_1' \\ (z_2'x)^{-1}z_2' \end{pmatrix} Au \\ &= Mu \quad \text{où} \quad M = (I - x(C'S^{-1}C)^{-1}C'S^{-1} \begin{pmatrix} (z_1'x)^{-1}z_1' \\ (z_2'x)^{-1}z_2' \end{pmatrix})A \end{aligned} \quad (53)$$

et $V(e_H) = E(e_H e_H') = \sigma_u^2 MM'$. L'estimation de $\hat{\theta}$ par les MCG est $\hat{\theta} = (x'(MM')^{-1}x)^{-1}x'(MM')^{-1}y$. Le problème maintenant est de trouver l'inverse de MM' car cette matrice est singulière. L'inverse généralisé de Moore-Penrose⁹ résout ce problème.

L'inverse généralisé de MM' est $(MM')^+$ et on le calcule comme suit :

$(MM')^+ = H D^{-1} H'(MM')'$ où D est une matrice diagonale de dimension $r \times r$ ($r = N - K - 1$) et les éléments de la diagonale sont les valeurs propres (différentes de zéro) de $(MM')'MM'$. H , de dimension $N \times r$, est une matrice de vecteurs propres correspondants aux r valeurs propres de $(MM')'MM'$ soit : $(MM')'MM'$ est de dimension $N \times N$ et de rang $N - K - 1$.

⁹ À noter que cette matrice est unique (Theil, 1971). Voir Jäckel (2002) pour des applications financières.

$$D = \begin{pmatrix} \lambda_1 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \lambda_{N-K-1} \end{pmatrix} \text{ et } H = \begin{pmatrix} m_{11} & \cdot & \cdot & m_{1r} \\ m_{21} & m_{22} & \cdot & m_{2r} \\ \cdot & \cdot & \cdot & \cdot \\ m_{N1} & m_{N2} & \cdot & m_{Nr} \end{pmatrix}. \quad (54)$$

Enfin, le test F habituel peut être utilisé pour tester $\theta = 0$. Ce test a $(K, N-2K-1)$ degrés de liberté (K restrictions et $2K+1$ paramètres à estimer). Il est à noter que si $N \rightarrow \infty$ alors $Mu \rightarrow u$. Par conséquent, pour les grands échantillons, les MCO peuvent être utilisés pour estimer θ .

4. Performances relatives des estimateurs

Cette section se divise en deux parties : a) trouver des éléments d'explication (l'explication est valable pour β_E qui est une extension de β_H) de la meilleure performance de β_H sur β_D et β_P . Dans le cas où il y a présence d'erreurs de mesure dans les variables explicatives, essayer de comprendre pourquoi β_H (et β_E extension de β_H) semble, encore une fois, avoir une meilleure performance que β_L (M.C.O); b) effectuer des expériences de Monte Carlo pour comparer β_H et β_E relativement à β_L .

4.1. Variables et modèles utilisés pour l'application économique : la simulation Monte Carlo¹⁰

On utilise un modèle simple reliant la consommation totale annuelle¹¹ aux variables suivantes:

i) revenu annuel du chef de famille (\tilde{X}_1); ii) âge du chef de famille (\tilde{X}_2); iii) revenu annuel des autres membres de la famille (\tilde{X}_3); iv) nombre de personnes-semaines formant la famille durant l'année (\tilde{X}_4).

Pour garder une certaine homogénéité de l'ensemble des observations, on a retenu les observations pour lesquelles le revenu total du ménage se situait entre 25 000 \$ et 55 000 \$. Les paramètres utilisés pour générer la variable dépendante sont :

$$y = 1,0 + 1,0\tilde{X}_1 + 1,0\tilde{X}_2 + 1,0\tilde{X}_3 + 1,0\tilde{X}_4 + u \quad (55)$$

¹⁰ Voir Jäckel (2002) et Glasserman (2004) pour des applications financières de cette méthode.

¹¹ Les données utilisées pour faire les expériences sont tirées de l'enquête faite par Statistique Canada sur les finances des Consommateurs.

où u est le terme d'erreur de régression. Les variables \tilde{X} observées dans l'échantillon sont supposées être les vraies valeurs des variables. Les paramètres ont été fixés à 1 dans le but de faciliter l'interprétation des résultats. Les Y ont donc été générés en utilisant l'équation ci-dessus avec un terme d'erreur normalement distribué ayant pour matrice de covariance $\sigma_u^2 I$. Dans les diverses expériences, σ_u a été fixé à différentes valeurs dans le but d'obtenir plusieurs valeurs du coefficient de détermination théorique, défini comme suit :

$$\tilde{R}^2 = (\beta \tilde{x}' \tilde{x} \beta) / (\beta \tilde{x}' \tilde{x} \beta + N \sigma_u^2) \quad (56)$$

où les \tilde{x}_i sont en déviations de la moyenne. Aussi, des erreurs de mesure normalement distribuées ont été ajoutées aux variables \tilde{X} . L'importance relative de ces erreurs de mesure a été évaluée pour chaque variable par le ratio de la variance des erreurs de mesure à la variance de la variable originale concernée $\lambda_i = \frac{\text{var}(v_i)}{\text{var}(\tilde{X}_i)}$. Des échantillons aléatoires de différentes

tailles ont été obtenus de l'ensemble total disponible de 4 400 observations. Dans toutes les expériences, les erreurs de mesure introduites affectent la variable \tilde{X}_1 (revenu annuel du chef de famille) ou bien les variables \tilde{X}_1 et \tilde{X}_3 (revenu annuel des autres membres de la famille). Pour mieux comprendre l'importance relative des erreurs de mesure affectant les variables explicatives dans les différentes expériences, il est utile de regarder le tableau 1. Par exemple, dans les expériences où seulement \tilde{X}_1 est affecté d'erreurs de mesure avec une variance de 5 % de la variance de \tilde{X}_1 , il est implicitement supposé que le revenu du chef de famille est rapporté avec une erreur moyenne absolue de 1 526 \$, ce qui correspond à 6,1 % de la valeur moyenne de \tilde{X}_1 . Il est aussi supposé que 95 % des répondants vont rapporter la valeur de \tilde{X}_1 avec une erreur absolue plus petite que 3 825 \$ (approximativement). Un cas d'erreur de 5 % est probablement une représentation plausible des erreurs de mesure présentes dans les enquêtes de revenu. Un pourcentage plus élevé correspondrait à des erreurs de mesure sur les variables qui sont plus difficiles à appréhender, comme la richesse personnelle nette ou bien d'autres types de variables économiques, comme les variables agrégées macroéconomiques.

Tableau 1
Statistiques sommaires pour l'échantillon complet de 4 400 données

Variables	\tilde{X}_1 (\$)	\tilde{X}_2 (Années)	\tilde{X}_3 (\$)	\tilde{X}_4 (Pers.-Sem.)
Valeur min.	-17 802	20	-15 000	52
Valeur max.	67 832	80	62 713	468
Ecart type	8 553	13,06	8 183	67,32
Moyenne	25 116	42,36	8 116	167,49

Taille	Valeur	% de la moyenne	Valeur	% de la moyenne	Valeur	% de la moyenne	Valeur	% de la moyenne	
Variance									
Des erreurs									
5 %	2 σ_v	3 825	15,2	5,8	13,8	3 659	45,1	30,1	17,9
	E(v)	1 526	6,1	2,3	8,9	1 460	17,9	12,0	7,2
10 %	2 σ_v	5 409	2,05	8,2	19,5	5 175	63,8	42,6	25,4
	E(v)	2 158	8,5	3,3	7,8	2 064	25,4	17,0	10,1
15 %	2 σ_v	6 625	26,4	10,1	23,9	6 338	78,1	52,2	31,1
	E(v)	2 643	10,5	4,0	9,5	2 528	31,2	20,8	12,4
25 %	2 σ_v	8 553	34,1	13,1	30,8	8 183	100,8	66,3	39,6
	E(v)	3 412	13,6	5,2	12,3	3 265	40,2	26,5	15,8
50 %	2 σ_v	12 095	48,2	14,3	33,8	8 964	110,4	73,7	44,0
	E(v)	4 825	19,2	5,7	13,5	3 576	44,1	29,4	17,6

4.2. Éléments d'explication de la meilleure performance de β_H par rapport à β_D et β_P

Dans le but d'expliquer la meilleure performance de β_H en regard de β_D et β_P dans le cadre de notre échantillon, il faut d'abord calculer la valeur relative de l'écart type des estimateurs β_D, β_P et β_H quand il n'y a pas d'erreur sur les variables explicatives, soit en divisant les écarts types de ces estimateurs par ceux des β_L correspondants (voir tableau 2).

Tableau 2
Écart type des estimateurs indiqués / écart type des MCO

Paramètres	β_D	β_P	β_H
β_1	10,4	4,6	3,6
β_2	9,9	18,5	2,0
β_3	4,0	5,6	2,1
β_4	25,3	6,2	2,4

On constate au tableau 2 le gain remarquable d'utiliser β_H au lieu de β_D et β_P . Par exemple, si on prend l'estimation du coefficient β_4 , les ratios de β_D et β_P sont respectivement de 25 et 6 pour 1 tandis que pour β_H le ratio est seulement de 2,4 pour 1. Cependant, il est important de noter que les résultats présentés dans ce tableau dépendent des données utilisées pour ces calculs, ceci n'étant qu'une illustration. La description de ces données a été présentée dans la section précédente.

Sachant que les trois estimateurs peuvent être interprétés comme des estimateurs à variables instrumentales obtenus en remplaçant les variables explicatives X par \hat{X} où \hat{X} est obtenu en régressant X sur z_1 ou sur z_2 ou bien sur les deux, et sachant que les variances de ces estimateurs dépendent de la corrélation de ces variables avec les variables explicatives, il peut être utile d'examiner la racine du coefficient de détermination de la régression des quatre variables \tilde{X}_i sur z_1 ou z_2 ou bien sur les deux (voir tableau 3).

Tableau 3
Coefficients de corrélation multiple de la régression de \tilde{X}_i sur z_i

Variables dépendantes	Variables indépendantes		
	z_1	z_2	z_1 et z_2
	Estimateurs correspondants		
	β_D	β_P	β_H
\tilde{X}_1	0,42	0,26	0,44
\tilde{X}_2	0,51	0,16	0,69
\tilde{X}_3	0,63	0,25	0,78
\tilde{X}_4	0,27	0,44	0,51

Pour mieux comprendre comment β_H peut avoir une meilleure performance que β_D et β_P , il est utile d'étudier, au tableau 3, la contribution marginale sur le $\sqrt{R^2}$ de l'ensemble des variables contenues dans z_2 une fois que l'ensemble des variables contenues dans z_1 a été introduit comme instrument pour \tilde{x} et aussi la contribution marginale au $\sqrt{R^2}$ de z_1 une fois que z_2 a été introduit comme instrument pour \tilde{x} . En d'autres termes, on essaie de voir si le fait d'ajouter z_1 à z_2 ou z_2 à z_1 contribue à augmenter le $\sqrt{R^2}$ en regard des contributions

idiosyncratiques de z_1 et z_2 . On constate que dans le cas de \tilde{X}_1 , \tilde{X}_2 et \tilde{X}_3 l'introduction de z_2 n'augmente pas fortement la corrélation multiple en comparaison de celle obtenue seulement avec z_1 (tableau 3). Cependant, le cas est différent pour z_2 .

4.3. Résultats : expériences de Monte Carlo¹²

Différentes mesures comme la racine carrée de l'écart quadratique moyen (REQM)¹³, l'erreur de type I, le biais, sont utilisées dans les différentes expériences de Monte Carlo servant à mesurer la performance relative de β_H et β_E par rapport à β_L .

La REQM mesure la dispersion de l'estimateur autour de la vraie valeur. Un bon estimateur en regard d'un moins bon a un REQM plus faible. On le définit comme suit¹⁴ :

$$REQM(\beta_L) = \sqrt{E[\beta_L - \beta]^2} = \sqrt{\text{biais}(\beta_L)^2 + \text{variance}(\beta_L)}. \quad (57)$$

Le biais mesure la différence entre les vraies valeurs d'un ensemble de paramètres et la moyenne des estimations de ces paramètres. Si cette différence est non nulle, alors on dit que l'estimateur est biaisé.

$$\text{biais}(\beta_L) = E(\beta_L) - \beta. \quad (58)$$

L'erreur de type I est la probabilité de rejeter l'hypothèse nulle (H0), qui suppose qu'un paramètre k est égal à une valeur donnée, alors que l'hypothèse est vraie. On la calcule de la façon suivante : $[1 - (\text{nombre fois qu'H0 est accepté} / \text{nombre d'échantillons})] \times 100$.

Pour avoir le nombre de fois que H0 est accepté, un test-t est effectué au niveau 5 % pour un paramètre k sur chaque échantillon.

Les différentes expériences de Monte Carlo concernent β_H et β_E relativement à β_L . La performance relative de ces estimateurs est mesurée en terme de REQM, de biais et d'erreurs de type I. La puissance du test est aussi discutée, étant définie comme le rapport entre le nombre de fois que H1 est acceptée alors que H1 est vraie et le nombre d'échantillons avec erreurs de mesure. Seulement les cas où les erreurs de mesure affectent une ou deux variables ont été considérés et dans ce dernier cas, on a supposé que les erreurs de mesure des deux variables ne sont pas corrélées. Il est à noter que les expériences ont été effectuées

¹²Voir l'appendice D pour le programme de la simulation Monte Carlo.

¹³ Il existe une extension de ce critère d'évaluation tenant compte de l'hétéroscédasticité nommée HRMSE. Voir Martens (2002) et Racicot (2003), chapitre 1.

¹⁴Voir l'appendice C pour la démonstration de cette relation.

en fixant la probabilité que les échantillons comportent des erreurs de mesure à 50 %. Il est probable que si les expériences avaient été effectuées avec 100 % des échantillons affectés d'erreurs, la performance des estimateurs proposés aurait augmenté relativement à celle de β_L (MCO) mais non la performance relative de β_E par rapport à β_H . D'autres éléments comme la matrice de corrélation des variables indépendantes ou la matrice de corrélation des erreurs de mesure peuvent aussi affecter la performance relative de β_H et β_E .

Pour évaluer les performances relatives de β_H et β_E , trois critères ont été retenus : le biais, la REQM et la probabilité d'erreur de type I lorsque l'on teste l'hypothèse (H0) que le paramètre estimé est égal à sa vraie valeur. Trois expériences ont été effectuées : une expérience avec erreurs de mesure sur \tilde{x}_1 seulement, en fonction de λ pour un \tilde{R}^2 donné et avec un échantillon de 500 observations; une expérience avec erreurs de mesure sur \tilde{x}_1 et \tilde{x}_3 ($\lambda = 5\%$) en fonction de la taille de l'échantillon pour un \tilde{R}^2 donné; et enfin une expérience avec erreurs sur \tilde{x}_1 seulement ($\lambda = 25\%$) en fonction de la taille de l'échantillon (chaque expérience a été faite avec 250 échantillons).

Le biais asymptotique pour les MCO (β_L) provient du fait qu'il y a présence d'erreurs de mesure sur les variables explicatives. Dans le cas où il n'y a qu'une seule variable explicative : $plim \beta_L = \beta/(1 + \lambda)$. En faisant varier λ , on peut visualiser l'évolution du biais.

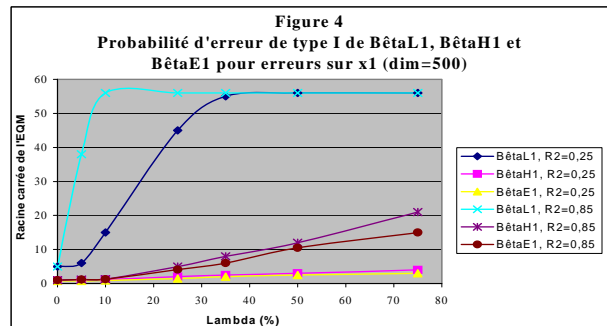
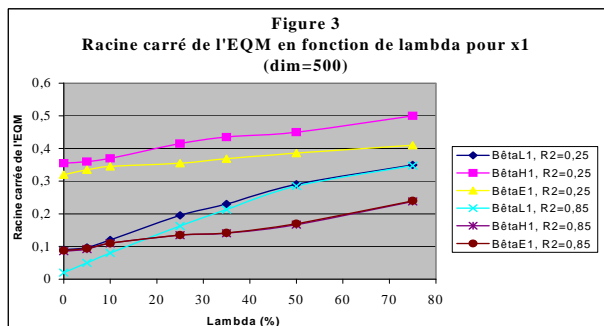
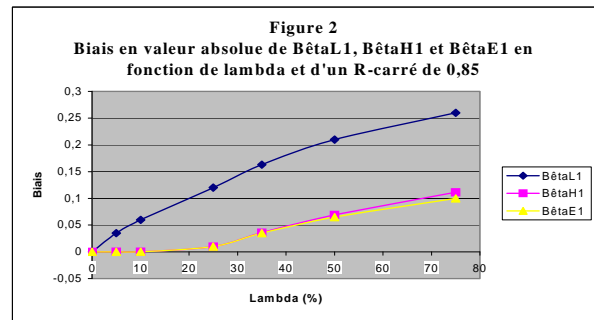
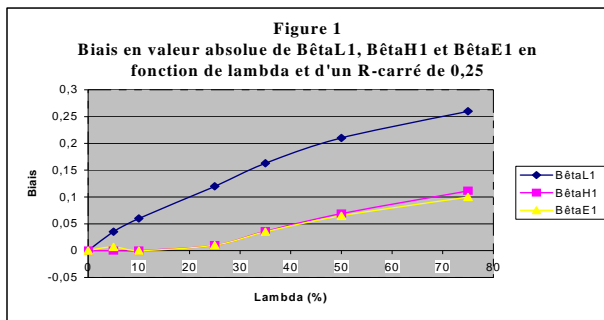
Les figures 1 et 2 présentent le biais des estimateurs $\beta_{L1}, \beta_{H1}, \beta_{E1}$ en fonction de λ pour un \tilde{R}^2 donné, pour un échantillon de 500 observations. Par exemple, pour $\tilde{R}^2 = 0,25$ (figure 1) si $\lambda = 10\%$, les biais respectifs de $\beta_{L1}, \beta_{H1}, \beta_{E1}$ sont environ de 0,06, 0,002, 0,01, ce qui veut dire : β_{H1} et β_{E1} ont une meilleure performance (pour ce critère) que β_{L1} . On constate aussi que β_{E1} ne performe pas mieux que β_{H1} pour tout λ . Si l'on reprend le même exemple, mais avec \tilde{R}^2 égal à 0,85 (figure 2), la performance de β_{H1} et β_{E1} demeure supérieure à celle de β_{L1} et cette fois β_{E1} surpasse β_{H1} pour tous les λ , mais de peu.

La figure 3 présente aussi les résultats de la première expérience mais cette fois-ci pour le deuxième critère retenu soit la REQM. On constate que les MCO (β_{L1}) pour un \tilde{R}^2 égal à 0,25 ont une meilleure performance que β_{H1} et β_{E1} , même si la taille des erreurs de mesure

est très grande. Par exemple, pour $\lambda = 75\%$, la REQM de β_{L_i} est d'environ 0,36 alors que celui de β_{H_i} et de β_{E_i} est d'environ 0,50 et 0,45. Maintenant si le \tilde{R}^2 égal à 0,85, on observe que β_{H_i} et β_{E_i} ont une meilleure performance que β_{L_i} à partir du moment où λ est supérieur à 10 %. On remarque aussi que β_{E_i} n'est toutefois pas plus performant que β_{H_i} .

La figure 4 présente le troisième critère retenu pour l'expérience, la probabilité d'erreur de type I, lorsqu'un test-t est effectué sur un paramètre. On remarque que la probabilité de commettre une telle erreur, lorsqu'un test est effectué sur un paramètre estimé par les M.C.O (β_L), est très élevée relativement à celle de β_{H_i} et de β_{E_i} . Par exemple, si $\tilde{R}^2 = 0,85$ et $\lambda = 10\%$ et que β_{L_i} est utilisé pour estimer le paramètre au lieu de β_{H_i} ou de β_{E_i} , la probabilité de commettre une erreur du type I est d'environ 56 % alors que cette probabilité est de seulement 2 % pour β_{H_i} et 1,2 % pour β_{E_i} .

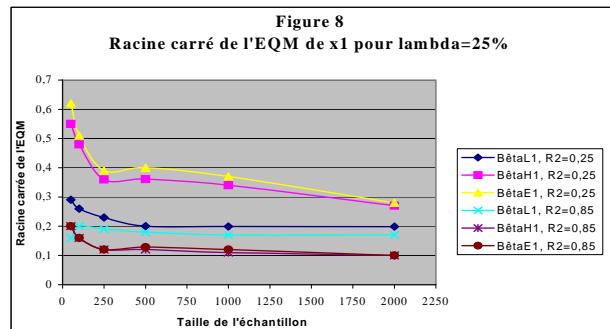
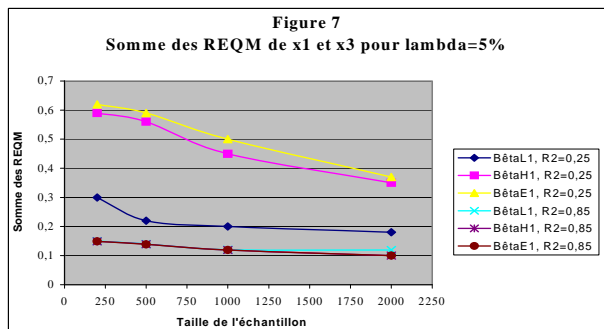
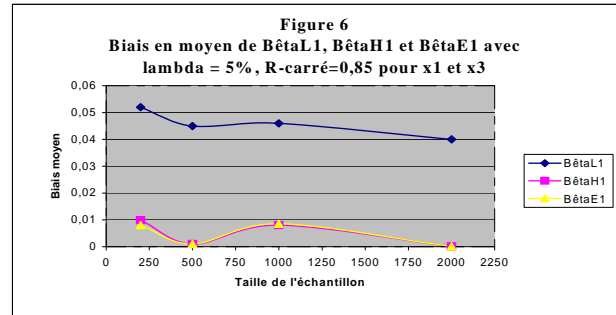
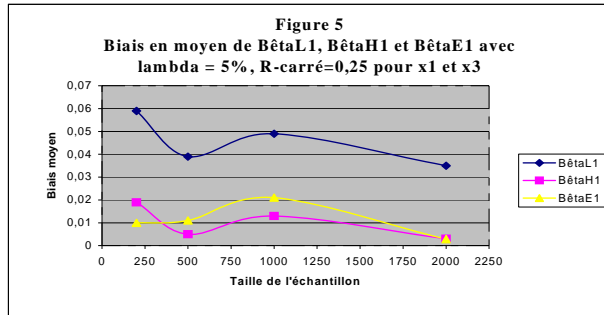
La deuxième expérience a été effectuée en introduisant des erreurs de mesure sur \tilde{x}_1 et \tilde{x}_3 d'une taille $\lambda = 5\%$. Les trois critères de performance des estimateurs sont évalués en fonction de la taille de l'échantillon.



Les figures 5 et 6 présentent le premier critère d'évaluation soit : les biais moyens de $\beta_{L_{13}}$, $\beta_{H_{13}}$ et $\beta_{E_{13}}$ en fonction de la taille de l'échantillon pour un \tilde{R}^2 donné (les indices 1 et 3 de $\beta_{L_{13}}$, $\beta_{H_{13}}$, $\beta_{E_{13}}$, signifient que c'est la moyenne des estimations des paramètres associés aux variables \tilde{x}_1 et \tilde{x}_3 qui est utilisée pour les différents critères d'évaluation). À la figure 5 ($\tilde{R}^2 = 0,25$), on remarque que le biais de $\beta_{L_{13}}$ est plus grand que celui de $\beta_{H_{13}}$ et de $\beta_{E_{13}}$ pour toute taille d'échantillons. Pour un échantillon de 2 000 observations, le biais de $\beta_{E_{13}}$ est plus petit que celui de $\beta_{H_{13}}$. Pour un \tilde{R}^2 égal à 0,85, les estimateurs sont encore plus performants car le biais, par exemple pour un échantillon de 2 000 observations, s'approche de zéro alors que celui de $\beta_{L_{13}}$ a augmenté par rapport à l'expérience faite avec le $\tilde{R}^2 = 0,25$. On voit aussi que $\beta_{E_{13}}$ semble être moins performant que $\beta_{H_{13}}$ pour l'ensemble des échantillons. La figure 7 montre la performance relative de $\beta_{L_{13}}$, $\beta_{H_{13}}$ et $\beta_{E_{13}}$ en terme de REQM en fonction de la taille de l'échantillon pour un \tilde{R}^2 donné. Lorsque le \tilde{R}^2 est égal à 0,25, on constate que la performance de $\beta_{L_{13}}$ est nettement supérieure à celle de $\beta_{H_{13}}$ et à celle de $\beta_{E_{13}}$ pour l'ensemble des échantillons testés. Dans le cas où \tilde{R}^2 est égal à 0,85, $\beta_{H_{13}}$ devient préférable à $\beta_{L_{13}}$ lorsque la taille de l'échantillon est supérieure à 1 000. Alors que $\beta_{E_{13}}$ surpasse $\beta_{L_{13}}$ seulement lorsque la taille de l'échantillon est supérieure à 1 200 observations.

La troisième expérience a été faite seulement pour la REQM des estimateurs β_{L_1} , β_{H_1} et β_{E_1} avec $\lambda = 25\%$ pour \tilde{x}_1 , en fonction de la taille de l'échantillon pour un \tilde{R}^2 donné. À la figure 8, on constate que β_{L_1} est supérieur à β_{H_1} et β_{E_1} avec $\tilde{R}^2 = 0,25$ pour les différents échantillons expérimentés. Mais on remarque aussi que l'écart entre les nouveaux estimateurs et β_{L_1} diminue à mesure que l'échantillon augmente. Pour \tilde{R}^2 égal à 0,85, β_{H_1} et β_{E_1} ont une performance supérieure à β_{L_1} lorsque l'échantillon a une taille supérieure à 75

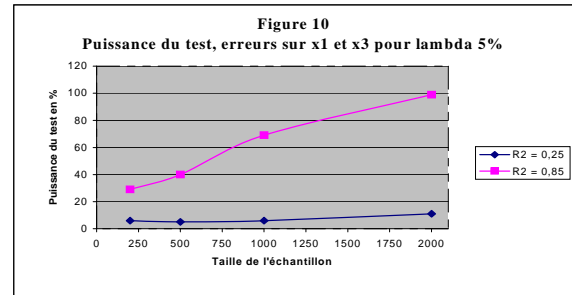
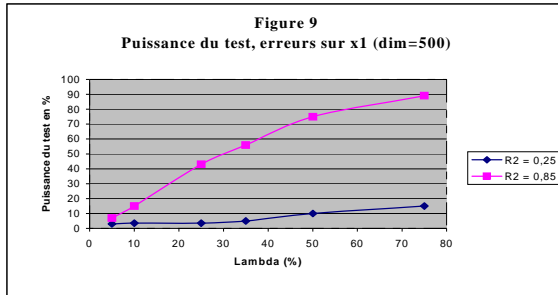
observations. La performance de β_{H1} semble supérieure à celle de β_{E1} pour l'ensemble des échantillons.



Deux expériences ont été faites pour évaluer la puissance du test d'erreurs de mesure sur les variables explicatives. La première a été effectuée avec un échantillon d'une taille égale à 500. La puissance du test est évaluée en fonction de λ pour un \tilde{R}^2 donné où seulement la variable \tilde{x}_1 est affectée d'erreurs de mesure. On peut voir les résultats de cette expérience à la figure 9. On remarque que le test est puissant seulement quand λ est très grand pour un \tilde{R}^2 de 0,85. Dans le cas où le \tilde{R}^2 est égal à 0,25, le test n'est pas puissant même pour des erreurs très grandes. Par exemple, si λ est égal à 75 %, la puissance du test n'est que d'environ 17 %.

La deuxième expérience effectuée avec des erreurs de mesure $\lambda = 5\%$ sur \tilde{x}_1 et \tilde{x}_3 est évaluée en fonction de la taille de l'échantillon pour un \tilde{R}^2 donné. À la figure 10, on constate que le test n'est pas très puissant lorsque le \tilde{R}^2 est égal à 0,25 même pour un échantillon d'une taille égale à 2000. Dans le cas où le \tilde{R}^2 est égal à 0,85, le test a une

puissance de 100 % quand l'échantillon compte 2 000 observations. Pour que la puissance du test dépasse 50 %, il faut que l'échantillon excède 700 observations.



Pour résumer les résultats des expériences avec les variables sur les finances consommateurs, on peut dire que :

- i) β_E ne semble pas tellement préférable à β_H pour les expériences effectuées. Les expériences effectuées supposent qu'environ la moitié des échantillons comportent des erreurs de mesure. La performance de β_E devrait s'améliorer si 100 % des échantillons comportent des erreurs de mesure.
- ii) β_H est sans biais lorsqu'il n'y a pas d'erreur de mesure. Donc quand il y a un doute sur la présence d'erreurs de mesure, il n'y a pas de risque sérieux d'utiliser β_H au lieu de β_L .
- iii) Quand il y a présence d'erreurs de mesure, β_H et β_E sont moins biaisés que β_L . Ceci demeure vrai même si la taille des erreurs de mesure (λ) est aussi petite que 5 %.
- iv) Lorsqu'il y a présence d'erreurs de mesure, la probabilité d'erreur du type I avec β_H et β_E semble toujours près de la valeur choisie (dans notre cas, la valeur choisie est de 5 %). Alors que si l'on utilise β_L , la probabilité d'erreur de type I demeure grande même si la taille des erreurs de mesure est aussi petite que $\lambda = 5$ %. β_E ne semble pas avoir une performance supérieure à β_H selon critère.
- v) La REQM de β_H et β_E est inférieure à celle de β_L pour des erreurs relativement faibles si le \tilde{R}^2 est élevé et l'échantillon suffisamment grand, lorsqu'une seule des variables est

affectée d'erreurs de mesure. Pour le cas où plus d'une variable est affectée d'erreurs de mesure, la performance relative des estimateurs β_H et β_E s'améliore.

Enfin, le test montre suffisamment de puissance si plus d'une variable explicative est affectée d'erreurs et si la taille de l'échantillon est relativement grande pour un \tilde{R}^2 élevé. Dans le cas où une seule des variables explicatives est affectée d'erreurs, ces erreurs doivent être grandes et le \tilde{R}^2 élevé pour que le test puisse les détecter. Donc le test n'est pas très puissant dans ce cas.

5. La méthode d'estimation à moments supérieurs : application au modèle du coût du capital de Benninga-Sarig¹⁵

Dans cette section, nous présentons la méthodologie et l'application de l'estimateur bêta H développé dans la section précédente au cas des erreurs sur les variables explicatives dans le cadre d'un modèle issu de la finance corporative¹⁶, soit le modèle du coût du capital avec taxes de Benninga-Sarig (1997).

5.1. Les données¹⁷

Dans cette application, nous utilisons les rendements de la compagnie Abbott Labs qui sont observés mensuellement. Ceux-ci sont calculés à partir des prix de clôture à la fin de chaque mois. Nous utilisons également comme variable explicative les rendements de l'indice boursier du Standard & Poor's 500 (S&P500) pour la même fréquence. Les rendements (sans dividendes) sont calculés à partir de la formule suivante :

$$\begin{aligned} y &= \frac{A.L._t - A.L._{t-1}}{A.L._{t-1}} \\ x &= \frac{SP500_t - SP500_{t-1}}{SP500_{t-1}} \end{aligned} \quad (59)$$

¹⁵Dans cet article, les estimateurs développés sont sous l'hypothèse de coupes transversales. L'auteur a dû faire certaines hypothèses de travail afin de généraliser les estimateurs aux séries temporelles pour effectuer la présente application. Il a également fait la conjecture que les statistiques t résultant de l'application d'un des estimateurs sont convergentes. Il s'agit d'une hypothèse simplificatrice permettant d'utiliser celle-ci afin de réaliser l'inférence statistique requise. Des clarifications à ce sujet seront apportées dans un article à venir.

¹⁶Voir Dagenais et Dagenais (1992a).

¹⁷D'autres applications de la méthode présentée dans ce texte sont possibles. En effet, les problèmes d'endogénéité et des erreurs de spécification ne se limitent pas qu'aux régressions classiques. Les données panels en sont un exemple. Racicot et Théoret (2007, 2008b) présentent une première ébauche de cette généralisation de la méthode développée dans cet article. Greene (2012) présente également un modèle en panel dynamique relié à la méthode de cet article (voir Anderson et Hsiao, 1981, 1982; Arellano et Bond, 1991; Arellano et Bover, 1995; Blundell et Bond, 1998).

où y et x sont, respectivement, la variable dépendante et explicative.

5.2. Méthodologie

Le modèle à estimer est celui de Benninga-Sarig (1997) qui est utilisé en finance corporative à des fins d'estimation du coût moyen du capital. Dans cette formulation, le taux de rendement (coût) des actions (coût de l'équité ou *cost of equity*, en anglais) se calcule comme suit :

$$r_{sL} = r_f(1 - t_c) + \beta_{sL}[E(r_m) - r_f(1 - t_c)] \quad (60)$$

r_{sL} : rendement (r) des actions (s : stock) avec levier (L)

r_f : taux d'intérêt sans risque (f : free)

$[E(r_m) - r_f]$: moyenne historique de la prime de risque du marché (m)¹⁸

β_{sL} : bêta des actions de la firme avec levier (risque financier).

Notons immédiatement que nous faisons ici la différence entre le bêta avec levier et celui sans levier. La théorie classique sur le sujet suggère que la relation entre ces deux quantités est la suivante :

$$\beta_{su} = \frac{\beta_{sL}}{[1 + (1 - t_c)L]} \quad (61)$$

Nous pouvons donc extraire l'information du bêta sans levier (risque d'affaires) à l'aide de données observées, c'est-à-dire le bêta avec levier obtenu des données du marché, le taux d'imposition des entreprises et le levier lui-même, L, défini comme étant le ratio de la dette (D) et de l'encours d'actions (S) (l'équité) : $L = D/S$. Cette quantité est nécessaire pour le calcul du coût d'opportunité du capital défini par :

$$\rho = r_f + \lambda\beta_{su} \quad (62)$$

où $\lambda = [E(r_m) - r_f]$ est la prime de risque du marché.

Ce modèle est également utilisé pour le calcul du coût de la dette qui est défini par :

$$r_{Dette} = r_f + \beta_{Dette}[E(r_m) - r_f(1 - t_c)] \quad (63)$$

¹⁸ Également définie comme étant le rendement excédentaire espéré du portefeuille du marché.

Étant donné les définitions du coût de l'équité (coût des actions) et du coût de la dette dans le cadre du modèle de Benninga-Sarig, nous pouvons maintenant effectuer le calcul du coût moyen du capital (WACC)¹⁹, une mesure apparaissant dans tous les états financiers des grandes entreprises :

$$WACC = \frac{S}{D+S} r_{SL} + \frac{D}{D+S} r_{Dette} (1-t_c) \quad (64)$$

où D peut représenter également l'encours d'obligations.

Comme il est couramment suggéré dans la littérature financière, et cela découle simplement du bon sens, il importe de comparer nos résultats à un autre modèle avant de pouvoir tirer des conclusions définitives. Un des modèles fréquemment utilisé dans cette littérature est celui de Gordon et ses extensions comme le modèle H²⁰ qui fait l'hypothèse d'un niveau d'abord anormalement élevé g_s de croissance du dividende et qui diminue par la suite à rythme constant égal à $2H$. Une fois cette période écoulée, le taux s'établit au niveau g_{long} . Nous n'irons pas plus loin sur ce sujet mais nous allons plutôt comparer nos résultats au modèle simple de Gordon que nous rappelons ici :

$$r_{SL} = \frac{d_0(1+g)}{P_0} = \frac{a RPA_0(1+g)}{P_0} + g = \frac{a(1+g)}{P_0/RPA_0} + g \quad (65)$$

où d_0 est le dernier dividende que la firme a payé, RPA_0 , le revenu par action au temps présent, P_0/RPA_0 , le rapport cours/bénéfice (en anglais, le P/E ratio de la firme), g , le taux de croissance (anticipé) des revenus de la firme et a , la proportion constante de ses revenus que la firme verse en dividendes. Cette formule a l'avantage d'utiliser des valeurs directement observables sur le marché.

Comme nous l'avons mentionné précédemment, le but de cette application est de comparer l'estimateur β_H à celui des MCO. L'estimateur β_E ne sera pas utilisé dans cette section car nous avons constaté que celui-ci n'était pas supérieur à β_H et cela pour toutes les expériences effectuées. De plus, β_E est relativement plus difficile à calculer que β_H . Pour toutes ces raisons, nous ne retiendrons que β_H que nous comparerons avec

¹⁹WACC : weighted average cost of capital. Voir Benninga (2000, 2008) ; Racicot et Théoret (2004). Un autre façon d'appeler le WACC est le taux d'actualisation ajusté pour le risque.

²⁰ $P_0 = \left[\frac{d_0(1+g_s)}{r-g_s} \right] \left[1 - \left(\frac{1+g_s}{1+r} \right)^N \right] + \frac{d_0(1+g_s)^N(1+g_{long}) + d_0(1+g_s)^N(H)(g_s-g_{long})}{(r-g_{long})(1+r)^N}$ où r , d_0 et N sont, respectivement, le taux

d'actualisation des dividendes, une annuité qui croît au taux g_s pendant N périodes. Voir Damodaran (1996, 2006); Fuller et Hsia (1984); Pinto et al. (2010); Racicot et Théoret (2004).

bêtaL (MCO) dans cette application concernant le calcul du coût moyen du capital (WACC). Au tableau 4 est présenté le programme EViews requis pour effectuer les calculs de l'estimateur bêtaH.

Tableau 4²¹

<pre>' Adaptation EViews de l'auteur au programme ' GAUSS présenté en appendice D. Notons que ' l'auteur a également généralisé cette application à ' d'autres cas en finance.</pre>	<pre>matrix(120,2) zz ' Construction de la matrice zz colplace(zz,z1,1) colplace(zz,z2,2)</pre>
<pre>' Estimateur BetaH pour les erreurs sur les variables ' Application de Beta H sur la SML et le modèle du ' coût du capital de Benninga-Sarig (modèle ajusté ' pour les taxes)</pre>	<pre>' Régression artificielle de px sur zz afin d'obtenir xchap ou ' hchap (wchap) ' ls px zz vector theta = (@inverse(@transpose(zz)*zz))*@transpose(zz)*px</pre>
<pre>' Calcul de z1 (zD de Durbin)</pre>	<pre>vector xchap=zz*theta vector wchap= px-xchap</pre>
<pre>genr x= sp5008998 genr r= abbot8998 genr xbarre=@mean(x) genr px1=x-xbarre genr z=px1^2 vector z1=z genr ybarre=@mean(r) genr py1=r-ybarre vector py=py1</pre>	<pre>matrix(120,2) ww colplace(ww,px,1) colplace(ww,wchap,2) ' Régression artificielle pour effectuer le test d'erreurs sur les ' variables et modèle corrigé pour celles-ci vector betah=(@inverse(@transpose(ww)*ww))*@transpose(ww)*py</pre>
<pre>' Calcul de z2 (ou zP de Pal)</pre>	<pre>mtos(px,px1) ' Transforme vecteurs en séries mtos(wchap,wchap1) mtos(py,py0)</pre>
<pre>genr w1=px1^3 vector w=w1 vector wt=@transpose(w) vector px=px1 vector pxt=@transpose(px) vector pxtt=pxt*px/@obs(px) vector diagon=@makediagonal(pxtt) ' Est un scalaire dans ce cas-ci vector z21=wt-(3*diagon*pxt/@obs(x)) vector z2=@transpose(z21)</pre>	<pre>' Résultat de la régression artificielle à comparer avec betaH ls py0 c px1 wchap1 ' Modèle de marché sans tenir compte des erreurs sur les variables ls r c x</pre>

On remarque rapidement la similitude entre ce programme et celui du langage de programmation du logiciel Gauss. On constate également la facilité avec laquelle on a pu implanter un tel estimateur dans un logiciel aussi courant et peu coûteux que EViews relativement aux progiciels tels Gauss ou Matlab. Comme nous l'avons expliqué précédemment, cet estimateur a l'avantage d'être très simple au niveau du calcul et ne requiert aucun choix, souvent difficile à faire, quant aux instruments. Tout comme les

²¹En 2005, l'auteur a découvert par essais et erreurs une nouvelle utilisation de ce code. En faisant une utilisation alternative du code pour en extraire les instruments et ensuite les insérer dans l'estimateur du GMM, un nouvel estimateur GMM a été obtenu (GMM-c et GMM-d). Voir Racicot et Théoret (2009, 2012).

estimateurs à variables instrumentales, il est convergent comme en témoignent nos nombreuses expériences de Monte Carlo. De plus, nous savons que cet estimateur a l'avantage d'être semi-paramétrique car il relève de la classe des estimateurs GMM. Finalement, cet estimateur est un GMM car nous savons que celui-ci recouvre les estimateurs à variables instrumentales.

5.3. Résultats

L'estimation du bêta entrant dans le modèle du coût du capital Benninga-Sarig effectuée à partir des rendements de Abbott Labs par la méthode des M.C.O et par la méthode d'estimation à moments supérieurs, soit notre nouvel estimateur, est présentée au tableau ci-dessous. Nous rappelons ici les deux modèles de marché estimés :

1) celui par les M.C.O, $\beta_L : r_t = \alpha + \beta r_{mt} + e_t$

2) celui par méthode à moments supérieurs, $\beta_H : r_t = \alpha + \beta r_{mt} + \beta_2 \hat{w}_{1t} + e_t$

Tableau 5

Estimation du bêta d'Abbott Lab à l'aide des MCO et de l'estimateur à moments supérieurs

	β_L	$\beta_H : PXI$	<i>WCHAPI</i>
Coefficients	0,8057	0,8822	-0,1494
t-Student	7,71	5,86	-0,71
R ²	0,3349	0,3379	-

On constate rapidement à la lecture du tableau 5 la régularité de l'estimation du modèle de marché. En effet, les R² des deux modèles sont similaires. On constate également une augmentation significative du bêta dans le modèle corrigé pour les erreurs sur les variables, c'est-à-dire modèle incluant le terme WCHAPI.

Passons maintenant au calcul du coût moyen du capital (WACC) à l'aide de l'estimation par les MCO du bêta, β_L , puis du β_H , cela en comparaison avec le modèle de Gordon. Pour ce faire nous avons besoin des informations suivantes²² pour la firme américaine Abbott Labs (1989-1998) :

²² Ces informations sont inspirées de l'ouvrage de Benninga (2000). Voir également Benninga (2008, 2011).

- i) taux sans risque (rendement bons du Trésor) : $r_f = 0,04$;
- ii) prime de risque (moyenne historique 89-98 du S&P 500) : $E(r_m) - r_f = (0,0132 * 12) - 0,04 = 0,1186$;
- iii) bêta d'Abbott Lab (voir tableau X) : $\beta_L = 0,8057$; $\beta_H = 0,8822$;
- iv) bêta de la dette²³ : $\beta_{\text{dette}} = 0,13$;
- v) taxe corporative : $t_c = 0,4$;
- vi) les coûts de l'équité ($r_{sL} = r_f(1-t_c) + \beta[E(r_m) - r_f(1-t_c)]$) et de la dette ($r_{\text{dette}} = r_f(1-t_c) + \beta_{\text{dette}}[E(r_m) - r_f(1-t_c)]$) sont:
- avec β_L : $r_{sL} = 0,04 * 0,6 + 0,8057 * (0,1584 - 0,04 * 0,6) = 0,1323$;
 - avec β_H : $r_{sL}^* = 0,04 * 0,6 + 0,8822 * (0,1584 - 0,04 * 0,6) = 0,1426$;
 - $r_{\text{dette}} = 0,04 * 0,6 + 0,13 * (0,1584 - 0,04 * 0,6) = 0,0415$;
- vii) l'encours des actions, S, et la dette de l'entreprise, D, sont :
- S = valeur au marché de l'équité (fin 98 en milliards\$) = 74,29
 - D = dette (en milliards\$) = 2,72.

En utilisant ces informations, nous pouvons comparer les WACC d'Abbott Labs pour notre nouvel estimateur et ceux du modèle de Benninga-Sarig (1997) :

$$1) \text{ WACC avec } \beta_L : WACC = \frac{S}{D+S} r_{sL} + \frac{D}{D+S} r_{\text{Dette}} (1-t_c) =$$

$$\frac{74,29}{74,29 + 2,72} 0,1323 + \frac{2,72}{74,29 + 2,72} 0,0415(1-0,4) = 0,1285$$

$$2) \text{ WACC avec } \beta_H : WACC = \frac{S}{D+S} r_{sL}^* + \frac{D}{D+S} r_{\text{Dette}} (1-t_c) =$$

$$\frac{74,29}{74,29 + 2,72} 0,1426 + \frac{2,72}{74,29 + 2,72} 0,0415(1-0,4) = 0,1376$$

²³ Étant donné qu'il n'est pas simple de calculer le bêta de la dette, nous l'estimons en utilisant les bêtas observés de bons avec différents niveaux de risque classifiés en termes de leurs échéances. Par exemple, on observe généralement qu'un bon de très courte échéance est classifié comme ayant un risque de bas niveau, soit un bêta égal à 0. Dans le cas extrême d'un bon de longue échéance (supérieure à 10 ans), le risque est généralement classifié comme étant intermédiaire, son bêta étant égal à 0,8. Dans le cas qui nous concerne, nous nous situons entre un bon de courte échéance et d'échéance moyenne (1-3ans et 3-10ans), les bêtas généralement observés pour ces intervalles étant de 0,1 et 0,35. Nous retenons donc pour notre exemple une approximation pour le bêta de la dette, à hauteur de 0,13, en comparaison avec ceux des bons.

Nous pouvons maintenant comparer ces résultats, comme il est d'usage en finance corporative, au résultat obtenu du modèle bien connu de Gordon. Pour ce faire, nous avons besoin des informations suivantes :

i) le ratio cours-bénéfice (une approximation) de l'entreprise (P/E ratio) :

$$P_0 / RPA_0 = 20;$$

ii) la proportion des revenus payés en dividendes (une estimation) : $a = 35\%$;

iii) la croissance des revenus (estimation) : $g = 6\%$.

En insérant ces informations dans la formule de Gordon, on obtient ainsi le coût de l'équité

$$r_{sL}^{**} = \frac{a * (1 + g)}{P_0 / RPA_0} + g = \frac{0,35(1,06)}{20} + 0,06 = 0,079$$

Le WACC, en utilisant le modèle de Gordon, est donc de :

$$WACC = \frac{S}{D + S} r_{sL}^{**} + \frac{D}{D + S} r_{Dette} (1 - t_c) = \frac{74,29}{74,29 + 2,72} 0,079 + \frac{2,72}{74,29 + 2,72} 0,0415(1 - 0,4) = 0,0771$$

Le tableau 6 présente un résumé de nos résultats.

Tableau 6

Coût moyen du capital (WACC) pour la firme Abbott Labs selon différents modèles et techniques d'estimation

Modèle de Gordon :	7,71%
Modèle de Benninga-Sarig avec β_L :	12,85%
Modèle de Benninga-Sarig avec β_H :	13,76%

À la lecture du tableau 6, on constate la grande différence qu'il existe entre le modèle de Benninga-Sarig et celui de Gordon et cela tant au niveau de notre nouvel estimateur ou de celui des MCO. En effet, la différence en pourcentage est respectivement de, pour le modèle de Benninga-Sarig avec MCO et celui utilisant notre nouvel estimateur, $(0,1285 - 0,0771) /$

0,0771 = 67%, et de, $(0,1376-0,0771) / 0,0771 = 78\%$. Cette grande différence entre les WACC obtenus des deux modèles est pour le moins peu rassurante. L'une des explications possibles, nonobstant les erreurs sur les variables causées par le portefeuille du marché puisque nous en avons tenu compte dans notre calcul, est la possibilité d'avoir utilisé des estimations peu représentatives de la réalité dans le modèle de Gordon. Par exemple, nous avons peut-être utilisé des valeurs sous-estimées pour les paramètres g et a . Nous avons peut-être aussi incorrectement estimé la valeur du P/E ratio. Une autre raison pourrait être que le modèle simple de Gordon est tout bonnement faux pour ce cas. Une extension de ce modèle comme celui du modèle H de Gordon pourrait probablement faire l'affaire. Pour ce qui concerne l'estimation du WACC de la firme Abbott Labs avec le modèle de Benninga-Sarig, on constate une différence appréciable mais dans un intervalle qui semble inclus dans un domaine plausible. En effet, la différence en pourcentage entre le modèle de Benninga-Sarig avec les MCO et celui utilisant notre nouvel estimateur est de $(0,1376-0,1285) / 0,1285 = 7,1\%$. Cette différence est justifiée par la présence probable d'erreurs sur les variables mais nous savons que notre estimateur est convergent même en présence de celles-ci, c'est-à-dire, entre autres, que le biais disparaît asymptotiquement, ce qui n'est certes pas le cas pour les MCO. Enfin, une dernière explication qui pourrait rendre compte de la si grande différence entre le modèle de Gordon et celui de Benninga-Sarig est l'éventualité que le CAPM ne soit tout simplement pas adéquat pour ces données²⁴. Mais notre application semble échapper à cette éventualité, ce qui nous conforte quant à la validité du CAPM et des estimations qui en résultent.

²⁴En effet, il existe des cas où il est reconnu que celui-ci ne tient tout simplement plus. Prenons par exemple le cas possible d'une firme risquée, c'est-à-dire que la variance de ses rendements est élevée, mais qui a un bêta négatif et un R^2 approximativement de 0.

6. Conclusion

La présence d'erreurs de mesure dans les variables explicatives est un problème majeur en analyse économétrique. Les expériences de Monte Carlo effectuées dans cet article ont montré que le fait d'ignorer cette réalité peut mener à des erreurs d'inférence importantes si les techniques d'estimation utilisées ne sont pas appropriées. Nous avons suggéré un nouvel estimateur β_H et son extension β_E . Cet estimateur est plus performant que l'estimateur des moindres carrés ordinaires (MCO) lorsqu'il y a présence d'erreurs de mesure dans les variables si ces erreurs sont relativement importantes. Nous avons aussi suggéré un test pour la présence d'erreurs dans les variables. Ce test s'est montré suffisamment puissant si plus d'une variable explicative est affectée d'erreurs et si ces erreurs sont relativement élevées. De plus, la puissance du test augmente quand le \tilde{R}^2 augmente.

D'autres expériences devraient être menées sur la puissance du test pour arriver à des conclusions définitives. Il est à noter que des expériences non décrites ici suggèrent que la corrélation entre les vraies variables indépendantes peut avoir un effet important sur la performance du nouvel estimateur (β_H). De plus, l'effet de la corrélation entre les erreurs de mesure des différentes variables n'a pas été présenté dans cet article.

En terminant, nous avons appliqué un de nos estimateurs, l'estimateur bêta H, à un cas financier. Dans ce cas, nous nous sommes intéressé à l'estimation du coût du capital, lequel utilise une estimation du bêta d'un actif financier. Nous avons également comparé ces résultats à l'estimateur des MCO que nous avons également comparé au modèle de Gordon. Les résultats obtenus sont encourageants. En effet, nous avons obtenu des résultats, pour notre première comparaison, qui vont dans le même sens que ceux obtenus dans la littérature financière qui a trait à l'estimation des bêtas (Blume, 1971, 1975; Broquet et al., 1997; Pinto et al., 2010). A savoir que les bêtas, une fois stabilisés, convergent vers 1. Nous voyons également nos estimations se comporter de la sorte. Notre estimateur semble donc, dans ce cas d'application, avoir un comportement qui encourage son utilisation. D'autant plus qu'il est facile d'implanter bêtaH dans tout logiciel favori de l'utilisateur. Notons enfin que nous avons illustré l'utilisation de notre nouvel estimateur dans le cadre d'un exemple issu de la finance corporative. Il va sans dire que celui-ci n'est pas limité à de telles applications. En effet, il peut être généralisé à toute situation empirique retrouvée autant en finance qu'en économie où la présence d'erreurs sur les variables est appréciable. Une généralisation

intéressante serait l'application de celui-ci aux données de panels. Celles-ci sont de plus en plus accessibles aux chercheurs tant en finance corporative, en finance bancaire qu'en marketing, par exemple. Ces champs de recherche pourraient bénéficier grandement de notre contribution.

Références

- Anderson, T. et C. Hsiao (1981), 'Estimation of dynamic models with error components', *Journal of the American Statistical Association*, 76, pp. 598-606.
- Anderson, T. et C. Hsiao (1982), 'Formulation and estimation of dynamic models using panel data', *Journal of Econometrics*, 18, pp. 67-82.
- Arellano, M. et S. Bond (1991), 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equation', *Review of Economic Studies*, 58, pp. 277-297.
- Arellano, M. et O. Bover (1995), 'Another look at the instrumental variables estimation of error component models', *Journal of Econometrics*, 68, pp. 29-52.
- Blundell, R. et S. Bond (1998), 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics*, 87, pp.115-143.
- Benninga, S. (2011), *Principles of Finance with Excel*, 2nd edition, Oxford, Oxford.
- Benninga, S. (2008), *Financial Modeling*, 3rd edition, MIT Press, Massachusetts.
- Benninga, S. (2000), *Financial Modeling*, 2nd edition, MIT Press, Massachusetts.
- Benninga, S. et O.H. Sarig (1997), *Corporate Finance : A Valuation Approach*, McGraw-Hill, New-York.
- Blume, M. (1971), 'On the assement of risk', *Journal of Finance*, 26, pp. 1-10.
- Blume M. (1975), 'Betas and their regression tendencies', *Journal of Finance*, 30, pp. 785-795.
- Broquet, C., Cobbaut, R., Gillet, R. et A. Berg (1997), *Gestion de Portefeuille*, 3^e édition, Collection Comptabilité, contrôle & finance, DeBoeck Université, Bruxelles.
- Coen, A. et F.E. Racicot (2007), 'Capital asset pricing models revisited: Evidence from errors in variables', *Economics Letters*, 95, pp. 443-450.
- Cochrane, D. et G. Orcutt (1949), 'Application to the least squares regression to relationships containing autocorrelated error terms', *Journal of the American Statistical Association*, 44, pp.32-61.
- Dagenais M.G. et D. L. Dagenais (1997), 'Higher moment estimators for linear regression models with errors in the variables', *Journal of Econometrics*, 76, pp. 193-221.
- Dagenais M.G. (1994), 'Parameter estimation in regression models with errors in the variables and autocorrelated disturbances', *Journal of Econometrics*, 64, pp. 145-163.
- Dagenais M.G. et D. L. Dagenais (1992a), 'Discriminating between consumption and market betas using a new estimator for regressions with errors in the variables', manuscrit, C.R.D.E., Université de Montréal.
- Dagenais M.G. et D. L. Dagenais (1991, 1992b), 'Estimation and testing in regression models with errors in the variables', manuscrit, C.R.D.E., Université de Montréal.
- Damodaran, A. (1996), *Investment Valuation*, Wiley.
- Damodaran, A. (2006), *Damodaran on Valuation*, 2nd edition, Wiley.
- Davidson, R. et J. MacKinnon (1993), *Estimation and Inference in Econometrics*, Oxford University Press, New York.
- Doan, M.P. et C.T. Lin (2012), 'On the robustness of higher-moment factors in explaining average expected returns: Evidence from Australia', *Research in Internation Business and Finance*, 26, pp. 67-78.
- Drion, E.F. (1951), 'Estimation of parameters of a straight line and of the variance of the variables, if they are both subject to error', *Indagationes Mathematicae*, 13, pp. 265-260.
- Durbin, J. (1954), 'Errors in variables', *International Statistical Review*, 22, pp. 23-32.

- Fomby T.B., Hill, C.R. et S.R. Johnson (1984), *Advanced Econometric Methods*, Springer-Verlag, New York.
- Fuller, R.J. et Hsia, C. (1984), A simplified common stock valuation model, *Financial Analyst Journal*, 40, pp. 49-56.
- Geary, R.C (1942), 'Inherent relations between random variables', *Proc. Royal Irish Academy*, 47, pp. 63-76.
- Glasserman, P. (2004), *Monte Carlo Methods in Financial Engineering*, Springer.
- Greene, W.H. (2012), *Econometric Analysis*, 7th edition, Pearson.
- Gregoriou, G.N., Hübner, G., Papageorgiou, N. et F.D. Rouah (2005), *Hedge Funds: Insights in Performance Measurement, Risk Analysis, and Portfolio Allocation*, Wiley.
- Gujarati, D.N. et D.C. Porter (2009), *Basic Econometrics*, 5th edition, McGraw-Hill Irwin.
- Hahn, J. et Hausman, J. (2003), 'Weak instruments: Diagnosis and cures in empirical economics', *American Economic Review*, 93, pp.118-125.
- Hansen, L. (1982), 'Large sample properties of the generalized method of moments estimators', *Econometrica*, 50, pp. 1029-1054.
- Hausman, J. (1978), 'Specification tests in econometrics', *Econometrica*, 46, pp. 1251-1271.
- Heij, C., de Boer, P., Franses, P.H., Kloek, T. et H.K. van Dijk (2004), *Econometric Methods with Applications in Business and Economics*, Oxford University Press.
- Hsiao, C. (1979), 'Measurement error in a dynamic simultaneous equations model with stationary disturbances', *Econometrica*, 47, pp. 475-494.
- Imhof, P.J. (1961), 'Computing the distribution of quadratic forms in normal variables', *Biometrika* 48, pp. 419-426.
- Jäckel, P. (2002), *Monte Carlo methods in finance*, Wiley.
- Johnston, J. et J. DiNardo (1997), *Econometric Methods*, 4th edition, McGraw-Hill, New York.
- Johnston, J. (1985, 1988), *Méthodes Économétriques*, 3^e édition, Tome 1 et 2, Économica, Paris.
- Judge G.G, C. Hill, W. Griffiths R., H. Lutkepohl et T. Lee (1985), *The Theory and Practice of Econometrics*, John Wiley & Sons, New York.
- Lambert, M. (2012), 'Hedge Fund Market Risk Exposures: A Survey', *Finance*, 33, pp. 39-78.
- MacKinnon, J. (1992), 'Model specification tests and artificial regressions', *Journal of Economic Literature*, pp. 102-146.
- Martens, M. (2002), 'Measuring and forecasting S&P 500 index-futures volatility using high-frequency data', *Journal of Futures Markets*, 22, pp. 497-518.
- Meng, J.G., Hu, G. et J. Bai (2011), 'Olive: A simple method for estimating betas when factors are measured with errors', *Journal of Financial Research*, 34, pp. 27-60.
- Morgenstern, O. (1972), *L'illusion statistique : précision et incertitude des données économiques*, Dunod, Paris.
- Pal, M. (1980), 'Consistent moment estimators of regression coefficients in the presence of errors in variables', *Journal of Econometrics*, 14, pp. 349-364.
- Pinto, J.E., Henry, E., Robinson, T.R. et Stowe, J.D. (2010), *Equity Asset Valuation*, CFA Institute, Wiley.
- Prais, S. et C. Winsten (1954), 'Trend estimation and serial correlation', *Cowles Commission Discussion Papers* No. 383, Chicago.
- Pindyck, R.S. et D. Rubinfeld (1998), *Econometric Models and Economic Forecasts*, 4th edition, Irwin McGraw-Hill.

- Racicot, F.E. (2003), 'Erreurs de mesure sur les variables économiques et financières', chapitre 3 de *Trois essais sur l'analyse des données économiques et financière*, Thèse de Ph.D. (Dir. Pr. Raymond Théoret), Université du Québec à Montréal.
- Racicot, F.E. (2000), 'Estimation et tests en présence d'erreurs de mesure sur les variables explicatives : vérification empirique par la méthode de simulation Monte Carlo', # 09-2000, École des sciences de la gestion, Université du Québec à Montréal (UQAM).
- Racicot, F.E. (1993), 'Techniques alternatives d'estimations et tests en présence d'erreurs de mesure sur les variables explicatives', essai de Maîtrise pour l'obtention du grade de M.Sc. en sciences économiques (économétrie) sous la direction du Pr. Marcel G. Dagenais, Département de sciences économiques, Université de Montréal.
- Racicot, F.E. et R. Théoret (2012), 'Optimally weighting higher-moment instruments to deal with measurement errors in financial return models', *Applied Financial Economics*, 22, pp. 1135-1146.
- Racicot, F.E. et R. Théoret (2009), 'On Optimal Instrumental Variables Generators, with an Application to Hedge Fund Returns', *International Advances in Economic Research*, 15, pp. 30-43.
- Racicot, F.E. et R. Théoret (2008b), *The Econometric Analysis of Hedge Fund Returns: An Errors-in-Variables Perspective*, Netbiblo, Spain.
- Racicot, F.E. et R. Théoret (2008a), 'On Comparing Hedge Fund Strategies Using New Hausman-based Estimators', *Journal of Derivatives & Hedge Funds*, 14, pp. 9-30.
- Racicot, F.E. et R. Théoret (2007) "The Beta Puzzle Revisited: A Panel Study of Hedge Fund Returns", *Journal of Derivatives & Hedge Funds*, 13, pp. 125-146.
- Racicot, F.E. et R. Théoret (2004), *Le calcul numérique en finance empirique et quantitative : ingénierie financière et Visual Basic*, 2^e édition, Presses de l'Université du Québec, Sainte-Foy.
- Statistique Canada (1988), Family Expenditure in Canada, 1986, Survey of Family Expenditure, Public Use Microdata file.
- Theil, H. (1971), *Principles of Econometrics*, John Wiley & Sons, New York.
- Theil, H. et A.S. Goldberger (1961), 'On pure and mixed statistical estimation in economics', *International Economic Review*, 2, pp. 65-78.
- White, H. (2001), *Asymptotic Theory for Econometricians*, Academic Press, Orlando.
- White, H. (1980), 'A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity', *Econometrica*, 48, pp. 817-838.

Appendice A

1. Preuve multivariée de la non-convergence des MCO

Soit :

$$y = \tilde{X} \beta + u$$

un modèle de régression linéaire multivariée sans terme constant, où \tilde{X} est une matrice de variables explicatives de dimension $N \times K$ avec $plim(\tilde{X}'\tilde{X}/N) = \Sigma_{\tilde{X}\tilde{X}}$, y est un vecteur $N \times 1$, et u un vecteur d'erreurs de régression $N \times 1$ avec $E(u) = 0$, $E(uu') = \sigma_u^2 I$. Supposons aussi que \tilde{X} n'est pas observé mais que X le soit où $X = \tilde{X} + v$ et v est une matrice d'erreurs de mesure sur les variables explicatives de dimension $N \times K$ avec $E(v) = 0$, $E(v_i, u) = 0$, $plim(v'v/N) = \Sigma_{vv}$.

Si l'on substitue \tilde{X} par sa valeur, on obtient :

$$\begin{aligned} y &= (X - v)\beta + u \\ &= \beta X - v\beta + u = \beta X + e \end{aligned}$$

où $e = u - v\beta$.

Si l'on applique les MCO sur $y = X\beta + e$ et qu'on remplace y par sa valeur, on obtient :

$$\hat{\beta} = (X'X)^{-1} X'X\beta + (X'X)^{-1} X'e = \beta + (X'X)^{-1} X'e.$$

Pour montrer que $\hat{\beta}$ n'est pas convergent lorsqu'il y a des erreurs de mesure sur les variables explicatives, il faut calculer le $plim\hat{\beta}$:

$$plim\hat{\beta} = \beta + (plim(X'X/N))^{-1} (plim(X'e/N))$$

où $plim(X'e/N) = plim(v'(u - v\beta)/N) = plim(v'u - v'v\beta/N) = -plim(v'v/N)\beta = -\Sigma_{vv}\beta$ et

$$plim(X'X/N) = plim[(\tilde{X}' + v')(\tilde{X} + v)/N] = plim[(\tilde{X}'\tilde{X}/N) + (\tilde{X}'v/N) + (v'\tilde{X}/N) + (v'v/N)] = \Sigma_{\tilde{X}\tilde{X}} + \Sigma_{vv}$$

$$plim\hat{\beta} = \beta - (\Sigma_{\tilde{X}\tilde{X}} + \Sigma_{vv})^{-1} \Sigma_{vv}\beta = \beta(I - \lambda) \text{ où } \lambda = (\Sigma_{vv} + \Sigma_{\tilde{X}\tilde{X}})^{-1} \Sigma_{vv}.$$

2. Solution au système d'équations du maximum de vraisemblance

On a les quatre équations suivantes :

$$1) \tilde{X}'(y - \tilde{X}'\tilde{\beta}) = 0$$

$$2) (X - \tilde{X}')'(X - \tilde{X}') = N\tilde{\sigma}_v^2$$

$$3) (y - \tilde{X}'\tilde{\beta})'(y - \tilde{X}'\tilde{\beta}) = N\tilde{\sigma}_e^2$$

$$4) (X - \tilde{X}')\tilde{\sigma}_e^2 + \tilde{\beta}'(y - \tilde{X}'\tilde{\beta})\tilde{\sigma}_v^2 = 0$$

Si l'on substitue $\tilde{\sigma}_e^2$ et $\tilde{\sigma}_v^2$ par leur valeur dans 4), on obtient :

$$(X - \tilde{X}') \frac{(y - \tilde{X}'\tilde{\beta})'(y - \tilde{X}'\tilde{\beta})}{N} + \tilde{\beta}'(y - \tilde{X}'\tilde{\beta}) \frac{(X - \tilde{X}')'(X - \tilde{X}')}{N} = 0.$$

En simplifiant, on obtient :

$$((y - \tilde{X}'\tilde{\beta})')' = (-\tilde{\beta}'(X - \tilde{X}')')' = y - \tilde{X}'\tilde{\beta} = -\tilde{\beta}'(X - \tilde{X}') = H$$

et si on multiplie 4) par $\tilde{\beta}$, on obtient 5) :

$$\tilde{\beta}'(X - \tilde{X}')\tilde{\sigma}_e^2 + \tilde{\beta}'(y - \tilde{X}'\tilde{\beta})\tilde{\sigma}_v^2 = 0.$$

Enfin, on n'a qu'à substituer dans 5) les coefficients de $\tilde{\sigma}_v^2, \tilde{\sigma}_e^2$ par H soit : $-H \tilde{\sigma}_v^2 + \tilde{\beta}^2 H \tilde{\sigma}_e^2 = 0$.
En simplifiant, on trouve $\tilde{\sigma}_v^2 = \tilde{\beta}^2 \tilde{\sigma}_e^2$.

Appendice B

Preuve de la convergence de β_H et de β_E

1. Pour montrer que β_H et β_E sont convergents, nous utilisons le concept de convergence en probabilité et de ce fait l'opérateur $plim$. Pour ce faire, il existe des théorèmes généraux qui sous des hypothèses appropriées permettent de calculer le $plim$ simplement en calculant l'espérance asymptotique, sans calculer la variance asymptotique, ce qui signifie qu'on n'a qu'à vérifier la nullité du biais asymptotique.

Il existe pour cela, entre autres, deux théorèmes (White, 2001):

Théorème 1: Observations indépendantes et distribuées de façon identique.

Si z_1, z_2, \dots, z_t est un ensemble de variables aléatoires et que $E(z_t) < \infty$ soit $E(z_t) = u$

et $cov(z_t, z_{t+n}) = 0$, alors $plim\left(\frac{\sum_{t=1}^N z_t}{N}\right) = \lim_{N \rightarrow \infty} E\left(\frac{\sum_{t=1}^N z_t}{N}\right)$.

Théorème 2: Observations indépendantes et distribuées de façon hétérogène.

Si z_1, z_2, \dots, z_t est un ensemble de variables aléatoires et que

$\exists \sigma > 0, 0 < \Delta < \infty / E(|z_t|^{1+\sigma}) < \Delta^{N \rightarrow \infty} < \infty, \forall t$, soit $E(z_t) = u_t$ alors,

$$plim \frac{\sum_{i=1}^N z_t}{N} = \lim_{N \rightarrow \infty} E \frac{\sum_{t=1}^N z_t}{N}.$$

2. Comme il a été mentionné précédemment, β_H et β_E sont convergents mais cela n'a pas été démontré. Nous prouvons maintenant que ces deux estimateurs sont effectivement convergents sous l'hypothèse H1. Pour cela, il faut montrer que β_D et β_P , sont convergents, ensuite l'on peut montrer que β_H et β_E le sont aussi. Sous l'hypothèse H1, on a:

$$X = \tilde{X} + v \text{ avec } V(\text{vech}(v)) = \Sigma \otimes I_N$$

$$y = \alpha i + \tilde{X}\beta + u$$

si on substitue \tilde{X} dans y , alors on a :

$$y = \alpha i + X\beta - v\beta + u$$

et en déviation de la moyenne, y devient :

$$y = x\beta + Au - Av\beta$$

donc pour β_D :

$$\beta_D = (z_1' x)^{-1} z_1' y$$

On suppose que $(z_1' x/N)^{-1}$ existe.

Si l'on substitue y dans β_D :

$$\beta_D = \beta + (z_1' x/N)^{-1} z_1' Au/N - (z_1' x/N)^{-1} (z_1' Av/N)\beta$$

Calcul du $plim \beta_D$:

$$plim \beta_D = \beta + [plim(z_1' x/N)]^{-1} plim(z_1' Au/N) - [plim(z_1' x/N)]^{-1} plim(z_1' Av/N)\beta$$

où

$plim(z_1' Au/N) = E_{N \rightarrow \infty}(z_1' Au/N) = 0$ parce que les éléments de z_1 et de Au sont indépendants et que $E(u) = 0$. Il nous reste à montrer que $plim(z_1' Av/N) = 0$ soit :

$plim(z_{1j}' A v_k / N)$ pour $(j, k = 1, \dots, K)$ et $\bar{v}_k = \frac{\sum_{i=1}^N v_{ik}}{N}$ alors,

$$plim \frac{\sum_{i=1}^N x_{ij}^2 (v_{ik} - \bar{v}_k)}{N} = plim(\sum x_{ij}^2 v_{ik} / N) - plim \bar{v}_k plim \sum x_{ij}^2 / N \text{ sachant que}$$

$plim \bar{v}_k = E_{N \rightarrow \infty}(\sum_{i=1}^N v_{ik}/N) = 0$ parce que $E(v_{ik}) = 0$ donc

$$plim \sum_{i=1}^N x_{ij}^2 (v_{ik} - \bar{v}_k)/N = plim \sum (\tilde{x}_{ij} + v_{ij})^2 v_{ik}/N$$

$$= plim \sum \tilde{x}_{ij}^2 v_{ik}/N + 2plim \sum \tilde{x}_{ij} v_{ij} v_{ik}/N + plim \sum v_{ij}^2 v_{ik}/N \text{ où}$$

- 1) $plim \sum (\tilde{x}_{ij}^2 v_{ik}/N) = E_{N \rightarrow \infty}(\tilde{x}_{ij}^2 v_{ik}/N) = 0$ car \tilde{x}_{ij}^2 et v_{ik} sont indépendants.
- 2) $plim \sum \tilde{x}_{ij} v_{ij} v_{ik}/N = E_{N \rightarrow \infty}(\sum \tilde{x}_{ij} v_{ij} v_{ik}/N) = \sum E(v_{ij} v_{ik}) E(\tilde{x}_{ij})/N = \mu_{jk}(1,1) \sum E(\tilde{x}_{ij})/N = 0$ car $E(\tilde{x}_{ij}) = E(\tilde{X}_{ij}) - \bar{X}_j = 0$ et $\mu_{jk}(1,1) = E(v_{ij} v_{ik})$ où $\mu_{ik}(l,m) = E v_{ij}^l v_{ik}^m$
- 3) $plim \sum v_{ij}^2 v_{ik}/N = E_{N \rightarrow \infty} \sum v_{ij}^2 v_{ik}/N = \sum \mu_{jk}(2,1)/N = 0$ où $\mu_{jk}(2,1) = E(v_{ij}^2 v_{ik}) = 0$ parce que les moments d'ordre impair de variables aléatoires normalement distribuées sont nuls (Theil, 1971).

Donc $plim \beta_D = \beta$.

Maintenant il faut montrer que β_p est convergent sous H1.

On a :

$$\beta_p = (z_2' x)^{-1} z_2' y$$

(on suppose que $(z_2' x)^{-1}$ existe) où $z_2' = z_3' - 3D(x'x/N)x'$ comme pour β_D , les éléments de β_p sont en déviation de la moyenne et donc $y = x\beta - Av\beta + Au$ est le modèle de régression sous H1 écrit en déviation de la moyenne. $\beta_p = \beta + (z_2' x)^{-1} z_2' Au - (z_2' x)^{-1} z_2' Av\beta$. Pour démontrer que β_p est convergent, il faut calculer le $plim \beta_p$ soit :

$$plim \beta_p = \beta + [plim(z_2' x/N)]^{-1} plim[(z_2' Au/N) - (z_2' Av\beta/N)].$$

β_p sera convergent si $plim [(z_2' Au/N) - (z_2' Av\beta/N)] = 0$.

En substituant z_2' par sa valeur dans cette équation, on obtient :

$$plim(z_3' Au/N) - plim(z_3' Av\beta/N) - 3D[plim(x'x/N)]plim(x' Au/N) + 3D[plim(x'x/N)]plim(x' Av\beta/N).$$

Sachant que u est indépendant de z_3 et de x , alors on a que :

$$plim(z_3' Au/N) = E_{N \rightarrow \infty}(z_3' Au/N) = 0 \text{ et que } plim(x' Au/N) = E_{N \rightarrow \infty}(x' Au/N) = 0 \text{ et}$$

$$plim(z_3' Av\beta/N) = plim(z_3' [I_N - ii'/N]v\beta/N) = plim(z_3' v\beta) - plim\left(\frac{z_3' [ii'/N]v\beta}{N}\right) \text{ où}$$

$$plim(z_3' [ii'/N]v\beta) = plim(z_3' i/N)plim(i'v\beta/N) = 0 \text{ et donc}$$

$$plim(z_3' Av\beta/N) = plim(z_3' v\beta/N)$$

De la même façon :

$$plim(x' Av\beta/N) = plim(x' v\beta/N).$$

Enfin, pour établir que $plim \beta_p = \beta$, il faut montrer que :

$$plim z_3' v/N + 3D(plim x'x/N) plim(x'v/N) = 0.$$

Pour un élément représentatif (j,k), on a par exemple :

$$plim(z_{3j}' v_k/N) = plim(\sum_{i=1}^N (\tilde{x}_{ij} + v_{ij} - \bar{v}_j)^3 v_{ik}/N)$$

$$= plim(\sum_{i=1}^N \tilde{x}_{ij}^3 v_{ik}/N) + plim(\sum_{i=1}^N v_{ij}^3 v_{ik}/N) - plim(\sum_{i=1}^N \bar{v}_j^3 v_{ik}/N)$$

$$+ 3plim(\sum_{i=1}^N \tilde{x}_{ij}^2 v_{ij} v_{ik}/N) - 3plim(\sum_{i=1}^N \tilde{x}_{ij}^2 \bar{v}_j v_{ik}/N) + 3plim(\sum_{i=1}^N \tilde{x}_{ij} v_{ij}^2 v_{ik}/N)$$

$$- 3plim(\sum_{i=1}^N v_{ij}^2 \bar{v}_j v_{ik}/N) + 3plim(\tilde{x}_{ij} \bar{v}_j^2 v_{ik}/N) + 3plim(\sum_{i=1}^N v_{ij} \bar{v}_j^2 v_{ik}/N)$$

$$- 6plim(\sum_{i=1}^N \tilde{x}_{ij} v_{ij} \bar{v}_j v_{ik}/N).$$

Le $plim(\sum_{i=1}^N \tilde{x}_{ij}^3 v_{ik}/N) = E_{N \rightarrow \infty}(\sum_{i=1}^N \tilde{x}_{ij}^3 v_{ik}/N) = 0$ parce que \tilde{x}_{ij}^3 et v_{ik} sont indépendants et que $E(v_{ik}) = 0$.

Le $plim(\sum_{i=1}^N v_{ij}^3 v_{ik} / N) = E_{N \rightarrow \infty}(\sum v_{ij}^3 v_{ik} / N) = \mu_{jk}(3,1) \neq 0$ car les moments pairs d'une distribution normale sont différents de zéro.

Le $plim(\sum_{i=1}^N \bar{v}_j^3 v_{ik} / N) = plim(\bar{v}_j^3 \sum_{i=1}^N v_{ik} / N) = plim \bar{v}_j \bullet plim(\sum v_{ik} / N)$ et donc :

$$plim(\sum_{i=1}^N v_{ik} / N) = E_{N \rightarrow \infty} \sum (v_{ik} / N) = E_{N \rightarrow \infty} \sum (E(v_{ik}) / N) = 0$$

$$plim(\sum_{i=1}^N \bar{v}_j^3 v_{ik} / N) = 0.$$

Le terme $3plim(\sum_{i=1}^N \tilde{x}_{ij}^2 v_{ij} v_{ik} / N) = 3 E_{N \rightarrow \infty}(\sum_{i=1}^N \tilde{x}_{ij}^2 v_{ij} v_{ik} / N) = \lim_{N \rightarrow \infty} 3(\sum \sigma_{\tilde{x}_j}^2 \mu_{jk}(1,1) / N) = 3 \sigma_{\tilde{x}_j}^2 \mu_{jk}(1,1).$

Le terme $3plim(\sum_{i=1}^N \tilde{x}_{ik} v_{ij}^2 v_{ik} / N) = -3plim(\sum_{i=1}^N (v_{ij}^2 \bar{v}_j v_{ik} / N) = 0$

parce que $E(v_{ij}^2 v_{ik}) = \mu_{jk}(2,1) = 0$ (les moments impairs d'une loi normale sont égaux à zéro).

Le terme :

$$3plim(\sum_{i=1}^N \tilde{x}_{ij} \bar{v}_j^2 v_{ik} / N) = 3plim(\bar{v}_j^2 \bullet \sum_{i=1}^N \tilde{x}_{ij} v_{ik} / N) = 3plim \bar{v}_j^2 \bullet plim(\sum_{i=1}^N \tilde{x}_{ij} v_{ik} / N)$$

et $plim(\sum_{i=1}^N \tilde{x}_{ij} v_{ik} / N) = E_{N \rightarrow \infty}(\sum_{i=1}^N \tilde{x}_{ij} v_{ik} / N) = 0$ car $E(v_{ik}) = 0 = E(\tilde{x}_{ij})$.

Le terme :

$$3plim(\sum_{i=1}^N v_{ij} \bar{v}_j^2 v_{ik} / N) = 3plim(\bar{v}_j^2) \bullet plim(\sum_{i=1}^N v_{ij} v_{ik} / N)$$

$$= 3plim(\bar{v}_j) \bullet plim(\bar{v}_j) \bullet plim(\sum_{i=1}^N v_{ij} v_{ik} / N) = 3 \times 0 \times 0 \times plim(\sum v_{ij} v_{ik} / N) = 0$$

car $plim(\bar{v}_j) = plim(\sum_{i=1}^N v_{ij} / N) = E_{N \rightarrow \infty}(\sum_{i=1}^N v_{ij} / N) = 0$ et $E(v_{ij}) = 0$.

Le terme :

$$-6plim(\sum_{i=1}^N \tilde{x}_{ij} v_{ij} \bar{v}_j v_{ik} / N) = -6plim(\bar{v}_j) \bullet plim(\sum_{i=1}^N v_{ij} v_{ik} / N) = 0$$

car $plim(\bar{v}_j) = 0$.

On a aussi pour un élément représentatif (j,k) que :

$$plim(x_j v_k / N) = plim(\sum_{i=1}^N (\tilde{x}_{ij} + v_{ij}) v_{ik} / N) = plim(\sum_{i=1}^N (\tilde{x}_{ij} v_{ik} + v_{ij} v_{ik} / N)$$

$$= E_{N \rightarrow \infty}(\sum_{i=1}^N (\tilde{x}_{ij} v_{ik} + v_{ij} v_{ik} / N) = \mu_{ik}(1,1)$$

Et que pour un terme représentatif (j,j) de la matrice diagonale $D(x'x/N)$:

$$plim(x_j' x_j / N) = plim(\sum_{i=1}^N (\tilde{x}_{ij} + v_{ij})^2 / N) = plim(\sum_{i=1}^N (\tilde{x}_{ij}^2 + 2 \tilde{x}_{ij} v_{ij} + v_{ij}^2) / N)$$

$$= plim(\sum_{i=1}^N \tilde{x}_{ij}^2 / N) + plim(\sum_{i=1}^N 2 v_{ij} \tilde{x}_{ij} / N) + plim(\sum_{i=1}^N v_{ij}^2 / N)$$

$$= E_{N \rightarrow \infty}(\sum_{i=1}^N \tilde{x}_{ij}^2 / N) + E_{N \rightarrow \infty}(2 \tilde{x}_{ij} v_{ij} / N) + E_{N \rightarrow \infty}(\sum_{i=1}^N v_{ij}^2 / N)$$

$$= \sigma_{\tilde{x}_j}^2 + 0 + \sigma_{v_j}^2$$

donc :

$$-plim(z_3' v / N) + 3Dplim(x'x / N) \bullet plim(x'u / N) =$$

$$= -\mu_{jk}(3,1) - 3 \sigma_{\tilde{x}_j}^2 \mu_{jk}(1,1) + 3(\sigma_{v_j}^2 + \sigma_{\tilde{x}_j}^2) \mu_{jk}(1,1)$$

$$= -\mu_{jk}(3,1) - 3 \sigma_{\tilde{x}_j}^2 \mu_{jk}(1,1) + 3 \sigma_{v_j}^2 \mu_{jk}(1,1) + 3 \sigma_{\tilde{x}_j}^2 \mu_{jk}(1,1)$$

$$= \mu_{jk}(3,1) + 3 \sigma_{v_j}^2 \mu_{jk}(1,1).$$

Sachant que les v_{ik} suivent une loi normale, et en utilisant les propriétés de la loi normale :

$$\mu_{jk}(3,1) = E(v_{ij}^3 v_{ik}) = 3E(v_{ij}^2) \bullet E(v_{ij} v_{ik}) = 3 \sigma_{v_j}^2 \bullet \mu_{jk}(1,1) \text{ donc } 3 \sigma_{v_j}^2 \mu_{jk}(1,1) - 3 \sigma_{v_j}^2 \mu_{jk}(1,1) = 0.$$

Enfin, pour montrer que β_H et β_E sont convergents,

$$\beta_H = (C' S^{-1} C)^{-1} C' S^{-1} \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} = W \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix}$$

$$\beta_E = (C' \hat{S}^{*-1} C)^{-1} C' \hat{S}^{*-1} \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix}$$

il faut calculer le $plim \beta_H$ et le $plim \beta_E$ soit :

$$plim \beta_H = [C'(plim(S/N))^{-1} C]^{-1} C' [plim(S/N)]^{-1} plim \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix}$$

$$\text{où } plim \begin{pmatrix} \beta_D \\ \beta_P \end{pmatrix} = \begin{pmatrix} plim \beta_D \\ plim \beta_P \end{pmatrix} = \begin{pmatrix} \beta \\ \beta \end{pmatrix} = C\beta \text{ et où } C = \begin{pmatrix} I_K \\ I_K \end{pmatrix} \text{ donc}$$

$$plim \beta_H = [C'(plim(S/N))^{-1} C]^{-1} C' [plim(S/N)]^{-1} C\beta = \beta \text{ et}$$

$$plim \beta_E = [C'(plim(\hat{S}^*/N))^{-1} C]^{-1} C' [plim(\hat{S}^*/N)]^{-1} C\beta = \beta . \text{ Q.E.D.}$$

3. On a supposé précédemment que les erreurs de régression et les erreurs de mesure sur les variables explicatives sont normalement distribuées. Cela peut sembler a priori une hypothèse peu raisonnable mais si on étudie plus en profondeur la composition des erreurs de mesure des variables explicatives et les erreurs de régression, alors il devient plus pertinent de faire de telles suppositions. Si on considère que :

$Y = \alpha_i + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u$ où u représente, par exemple, les variables explicatives omises et si on a omis $u = \beta_4 X_4 + \dots + \beta_N X_N$, où u est une somme de facteurs omis, chacun des facteurs n'étant pas nécessairement normalement distribué. La somme des facteurs, qui devrait tendre vers l'infini par le théorème central limite, convergera en distribution vers une loi normale²⁵ :

$$N^{-1/2} \sum_{i=1}^N (X_i \beta_i - E(X_i \beta_i)) \xrightarrow{d} N(0, \sigma_u^2 I).$$

Maintenant, pour les erreurs de mesure sur les variables explicatives, l'explication est la même en ce qui concerne l'intuition sous-jacente. Les v_i sont des sommes de facteurs omis et par le théorème central limite chaque vecteur v_i est normalement distribué, ce qui implique que l'ensemble des vecteurs v_i est normalement distribué (Johnston, 1985, 1988; Johnston et Dinardo, 1997; Judge et al., 1985; Gujarati, 2009).

4. On calcule la "p-value" associée à m comme suit :

$$Pr(\tilde{m} > m) = Pr[\eta' [J G^{-1} J - m [I - X(X'X)^{-1} X'] / (N - K - 1)] \eta] > 0$$

où $\tilde{m} = (u' J' G^{-1} J u / \sigma^2) / (\hat{\sigma}^2 / \sigma^2)$. Alternativement \tilde{m} peut être vu comme :

$\tilde{m} = ((\eta' J' G^{-1} J \eta) / [\eta' (I - X(X'X)^{-1} X') \eta] / (N - K - 1))$ où e_L est exprimé en terme de u et η . η est une variable normale réduite : $\eta = u / \sigma$.

Enfin, cette probabilité $Pr(\tilde{m} > m)$ peut se calculer par la procédure d'Imhof (Imhof, 1961).

²⁵ Pour la question de convergence, on a besoin d'un théorème plus général. D'après la version généralisée du théorème central limite (c.-à-d. le théorème central limite multivarié de Lindeberg-Feller) : si v_1, v_2, \dots, v_N sont des vecteurs aléatoires distribués indépendamment avec pour vecteurs de moyennes $\mu_1, \mu_2, \dots, \mu_N$ et de matrices de covariances

V_1, V_2, \dots, V_N et si les moments de troisième ordre existent : $E(v_{ij,N}^3) < \infty$, alors : $N^{-1/2} \sum_{i=1}^N (v_i - \mu_i) \xrightarrow{d} N(0, \bar{\Sigma})$ où $\bar{\Sigma} = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N V_i$ (Fomby et al., 1984; White, 2001; Greene, 2012).

Appendice C

Pour un élément i du vecteur $\hat{\beta}$ soit :

$$EQM(\hat{\beta}_i) = \text{Var}(\hat{\beta}_i) + \text{Biais}^2(\hat{\beta}_i)$$

Démonstration :

$$\begin{aligned} EQM(\hat{\beta}_i) &= E[\hat{\beta}_i - \beta_i]^2 = E[\hat{\beta}_i - E(\hat{\beta}_i) + E(\hat{\beta}_i) - \beta_i]^2 \\ &= E[\hat{\beta}_i^2 - \hat{\beta}_i E(\hat{\beta}_i) + \hat{\beta}_i E(\hat{\beta}_i) - \hat{\beta}_i \beta_i - E(\hat{\beta}_i) \hat{\beta}_i + E(\hat{\beta}_i)^2 - E(\hat{\beta}_i) E(\hat{\beta}_i) + \beta_i E(\hat{\beta}_i) \\ &\quad + \hat{\beta}_i E(\hat{\beta}_i) - E(\hat{\beta}_i)^2 + E(\hat{\beta}_i)^2 - E(\hat{\beta}_i) \beta_i - \beta_i \hat{\beta}_i + E(\hat{\beta}_i) \beta_i - \beta_i E(\hat{\beta}_i) + \beta_i^2] . \end{aligned}$$

En simplifiant, on obtient:

$$\begin{aligned} &E[\hat{\beta}_i^2 - 2\hat{\beta}_i E(\hat{\beta}_i) + E(\hat{\beta}_i)^2 - 2E(\hat{\beta}_i) \beta_i + E(\hat{\beta}_i)^2 + \beta_i^2] \\ &= E[(\hat{\beta}_i - E(\hat{\beta}_i))^2 + (\beta_i - E(\hat{\beta}_i))^2] \\ &= \text{Var}(\hat{\beta}_i) + \text{Biais}^2(\hat{\beta}_i) . \text{ Q.E.D.} \end{aligned}$$

Appendice D

Tiré de Racicot (1993) : essai en économétrie pour l'obtention du grade de M.Sc. en sciences économiques (économétrie), sous la direction du professeur Marcel G. Dagenais du département de sciences économiques de l'Université de Montréal.

Expériences de Monte Carlo pour β_H et β_E dans le progiciel GAUSS

```
new;
ajour1=date;
et = hsec;

/* LES PARAMÈTRES DE DÉPART À SPECIFIER SONT
INDIQUÉS PAR LES INDICATIONS QUI SE
TROUVENT À GAUCHE */

ater=6;
itera=1; /* Nombre de passes: itera */
caprim=4; /* Nombre de variables excluant la constante:
caprim */
ca=caprim+1;
let zrcarre= { .25 .85}; /* Rcarre .25 et .85 */

let zproporc= { 5 10 25 35 50 75,
00 0 0 0 0,
00 0 0 0 0,
00 0 0 0 0};

/* Proportion d'erreurs. Matrice: (caprim x
itera)*/

let zechan= {40 100 250 500 1000 2000}; /* Taille des
échantillons */

properr= 50; /* pourcentage d'échantillons avec erreurs de
mesure */

ztou=1;
output file=betahe.out reset; /* fichier d'output */

do while ztou<=ater;

/* do while ztourz<=itera; */

ztourz=1;

ztour=1;
iteraz=2;
do while ztour<=iteraz;

" "; " "; "numero du tour="; ztou; " ";
pourtage=0;

let seed2 = 612 612 612 612 612 612 ;
/* Mettre les "seed" ou 0. Vecteur: (itera) */

/* ATTENTION 1 */

dimen=zechan[.,ztou]; /* Dimension de l'échantillon */
nvec=250; /* Nombre d'échantillons */
niconH0 = 0.05; /* Niveau de confiance pour H0: pas d'erreur de
mesure */

nivconf=0.05; /* Niveau de confiance pour les erreurs de
type 1 */
rcarrez=zrcarre[.,ztour]; /* Rcarre */
proporc =zproporc[.,ztourz];

/* ATTENTION 2 */
/*let retoil[2,2] = 1 .5
.5 1; */

let retoil= 1 0 1 0 0 1 0 0 0 1;
/* Matrice de Correlation voulue pour les x */

/* ATTENTION 3 */
/*let rvetoil[2,2] = 1 0
0 1; */

let rvetoil= 1 0 1 0 0 1 0 0 0 1;
/* Matrice de Correlation voulue pour les "v"
*/

fxcorx=0; /* 0 si on veut garder les x originaux, 1
autrement */

/* ATTENTION 4 */
ff = "a:var11.dat"; /* Banque de donnees */

/* ATTENTION 5 */
retoil=xpnd(retoil);
deterx=det(retoil);
if deterx<=0;
" matrice de correlation des x non symmetrique definie positive";
stop;endif;

/* ATTENTION 6 */
rvetoil=xpnd(rvetoil);
deterv=det(rvetoil);
if deterv<=0;
"matrice de corrélation des v non symétrique définie positive";
stop; endif;

/* ATTENTION 7 */
open h=^ff; /* BETA36.DAT; */
xOR1=readr(h,dimen);
h1=close(h);
x1 = xOR1[.,1:(ca-1)];

/* CALCUL DE CERTAINS PARAMETRES */

format /m1/ro 12,7;
vb = ones(ca,1);
vbsc = vb[2:ca];
caprim2 = 2*caprim;
ivec = ones(dimen,1);
iivec = ones(caprim,1);
if seed2[ztourz] == 0;
```

```

seed2[ztourz] = rndu(1,1);
seed2[ztourz]=seed2[ztourz]*((2^31)-1);
seed2[ztourz] = round(seed2[ztourz]);endif;
propor = propor./100;
rx = corrx(x1);
{va1,ve1} = eigrs2(rx);
{va2,ve2} = eigrs2(retoil);
{va3,ve3} = eigrs2(rvetoil);
va1 = sqrt(va1);
va2 = sqrt(va2);
va3 = sqrt(va3);
va1a = zeros(caprim,caprim); va1a = diagrv(va1a,va1);
va2a = zeros(caprim,caprim); va2a = diagrv(va2a,va2);
va3a = zeros(caprim,caprim); va3a = diagrv(va3a,va3);
p = va1a*ve1';
petoil = va2a*ve2';
pvetoil = va3a*ve3';
if fxcorx==0; goto labelnou; endif;
t = inv(p) * petoil;
moyenx = meanc(x1);
xdev = x1 - moyenx';
ecartex = (sqrt(sumc(xdev.*xdev)/dimen));
normx = xdev./ecartex;
normmetol = normx * t * fxcorx;
xetoil = normmetol.*ecartex;
xetoil = xetoil + moyenx';
x1 = xetoil + moyenx'; clear xetoil;clear normmetol;
labelnou:
vx =ivec~X1;
coretoil = corrx(x1);

/* IMPRESSION DES PARAMETRES */

"PROGRAMME = BETAHE.PRG";
"PARAMETRES:" vb;
"CORRELATION DE X1=" rx;
"FXCORX=" FXCORX;
"CORRELATION DES X DESIREE SI FXCORX=1" retoil;
"CORRELATION DES X TRANSFORMES=" coretoil;
"DATE=" aujour1;
"DIMENSION=" dimen;

/* ATTENTION 8 */
"BANQUE DE DONNEES=" $ff;
"NVEC=" nvec;
"NIVEAU DE CONFIANCE DU TEST D'ERREURS SUR
VARIABLES=" NICONH0;
"NIVEAU DE CONFIANCE DU TEST D'ERREURS DE TYPE 1
=" nivconf;
"K=" ca;
format /m1/ro 20,15; "SEED2=" seed2[ztourz]; format /m1/ro
11,5;
"R CARRE" rcarrez;
"POURCENTAGE D'ERREUR=" propor;
"PROBABILITE D'ECHANTILLONS AVEC ERREURS DE
MESURE, EN %" PROPERR;

/* MISE A ZERO DE PARAMETRES QUI SERVIRONT POUR
LES CALCULS DES
MOYENNES ET VARIANCES DES ESTIMATEURS ET AUSSI
LES ACCEPTATIONS */

betaLm =zeros(ca,1); betaDm =zeros(ca,1); betaPm =zeros(ca,1);
betaLv =zeros(ca,1); betaDv =zeros(ca,1); betaPv =zeros(ca,1);
betaLv =zeros(ca,1); betaDv =zeros(ca,1); betaPv =zeros(ca,1);
betaHm =zeros(ca,1); betaHv =zeros(ca,1); betaHvv =zeros(ca,1);
betaEm =zeros(ca,1); betaEv =zeros(ca,1); betaEv =zeros(ca,1);

```

```

bPTHm =zeros(ca,1); bPTHv =zeros(ca,1); bPTHvv
=zeros(ca,1);
bPTEm =zeros(ca,1); bPTEv =zeros(ca,1); bPTEvv
=zeros(ca,1);
acephoH =zeros(1,1); x =zeros(dimen,caprim);
sVbLh0 =zeros(ca,1);
sVbDh0 =zeros(ca,1); sVbPh0 =zeros(ca,1);
sVbHh0 =zeros(ca,1);
sVbDh1 =zeros(ca,1); sVbPh1 =zeros(ca,1);
sVbHh1 =zeros(ca,1);
sVbEh0 =zeros(ca,1); sVbEh1 =zeros(ca,1);
svbpth =zeros(ca,1); svbpte =zeros(ca,1);
sEbLh0 =zeros(ca,1);
sEbDh0 =zeros(ca,1); sEbPh0 =zeros(ca,1);
sEbHh0 =zeros(ca,1);
sEbDh1 =zeros(ca,1); sEbPh1 =zeros(ca,1);
sEbHh1 =zeros(ca,1);
sEbEh0 =zeros(ca,1); sEbEh1 =zeros(ca,1);
sebpth=zeros(ca,1); sebpte=zeros(ca,1);
suctbLh0=zeros(ca,1); suctbDh0=zeros(ca,1);
suctbPh0=zeros(ca,1);
suctbHh0=zeros(ca,1); suctbEh0=zeros(ca,1);
suctbLh1=zeros(ca,1); suctbDh1=zeros(ca,1);
suctbPh1=zeros(ca,1);
suctbHh1=zeros(ca,1); suctbEh1=zeros(ca,1);
suctbpth=zeros(ca,1); suctbpte=zeros(ca,1);

/* CREATION DE NVEC VECTEURS DE DIMEN VALEURS
POUR
LA VARIABLE CONSOMMATION */

N = dimen;
r1 = 1;
c = zeros(caprim,caprim); cc = diagrv(c,ivec); C=cc/cc;
ccnou=eye(ca); cnou= ccnou/ccnou;
xbarre = ivec*x1/N; px = x1-ivec*xbarre;
standx = stdc(px); sigma2x = standx.*standx;
sigma2v = sigma2x.*propor;
sigmav = sqrt(sigma2v);
sigma2 = (vbsc'*px'*px*vbsc*(1-rcarrez)/(rcarrez*dimen));
sigma = sqrt(sigma2);
"SIGMA =" sigma;
"SIGMAV =" sigmav;
clear iivec;

/* DEBUT DE LA BOUCLE DO */

seed1=seed2[ztourz];
bonchoix=0;
do while r1<=nvec;
if r1==1;
bonde=1;
format /m1/ro 20,11; "seed1=" seed1; format /m1/ro 11,5; endif;
e1 = rndns(dimen,1,seed1);
e1 = sigma.*e1;
Y = vx*vb+e1; /* Calcul des vecteurs de consommation*/

e1 = rndns(dimen,caprim,seed1);
e1 = e1*pvetoil;
e1 = sigmav'.*e1;
if properr==100; goto labeldd; endif;
decid=rndus(1,1,seed1);
if decid>(properr/100); bonde=0; X=X1; goto labelff; endif;
labeldd:
pourtage=pourtage+1;
X = X1+e1;
labelff:

```

```

clear e1;

/* CALCUL DES BETAS, DES Q ET DES H */

xbarre = ivec*X/N; px = X-ivec*xbarre;
ybarre = ivec*Y/N; py = Y-ivec*ybarre;
x=ivec~x;

qL = inv(px*px)*px';
qLnou= inv(x*x)*x';
betaL = inv(px*px)*px*py;

erols=py-px*betaL;
betaLco=ybarre-xbarre*betaL;
betaL=betaLco|betaL;

betaLm = betaLm+betaL;
betaLv = betaL.*betaL;
betaLv = betaLv+betaLv;

zD = px.*px;
zDpa = zD'-(zD*ivec*ivec'/N);
qD = inv(zD*px)*zDpa;
zDnou= ivec~zD;
qDnou=inv(zDnou*x)*zDnou';
betaD = inv(zD*px)*zD*py;

netaD = py - px*betaD;
betaDco=ybarre-xbarre*betaD;
betaD=betaDco|betaD;

betaDm = betaDm+betaD;
betaDvv = betaD.*betaD;
betaDv = betaDv+betaDvv;

w = zD.*px;
mocro = px*px; diagon=mocro.*cc;
zP = (w'-(3*diagon*px'/n))';
zPnou= ivec~zP;
zPpa = zP'-(zP*ivec*ivec'/N);
qP = inv(zP*px)*zPpa;
qPnou=inv(zPnou*x)*zPnou';
betaP = inv(zP*px)*zP*py;

netaP = py - px*betaP;
betaPco=ybarre-xbarre*betaP;
betaP=betaPco|betaP;

clear w,mocro,diagon;
betaPm = betaPm+betaP;
betaPvv = betaP.*betaP;
betaPv = betaPv+betaPvv;

zz = zDpa'~zPpa';
theta = inv(zz*zz)*zz*px;
xcha = zz*theta;
wcha = px-xcha;
ww = px~wcha;
phi = inv(ww*ww)*ww*py;
betaH = phi[1:caprim];

netaH = py - px*betaH;
betaHco=ybarre-xbarre*betaH;
betaH=betaHco|betaH;

epe = erols*erols;
clear erols;
betaHm = betaHm+betaH;
betaHvv = betaH.*betaH;
betaHv = betaHv+betaHvv;

u=1; v1 = zeros(caprim,1); Do while u <=caprim; v1[u]=caprim+u;
u=u+1; endo;
www1 = inv(ww*ww);
gama = phi[ca:caprim2];
a22 = submat(www1,v1,v1);
psi = py-(ww*phi);
fft = ((gama*inv(a22)*gama)/caprim)/((psi*psi)/(dimen-
(2*caprim)-1));
ftoil = cdfc(fft,caprim,(dimen-(2*caprim)-1));
if niconH0<ftoil; acephoH = acephoH + 1; endif;
czero = zeros(ca,ca);
zdc=ivec~zd;
AA1 = inv((zDc*x)/dimen);
AA2 = inv((x*zDc)/dimen);
zpc=ivec~zpc;
BB1 = inv((zPc*x)/dimen);
BB2 = inv((x*zPc)/dimen);
CC1 = (AA1~czero)|(czero~BB1);
zDi = zDc;
zPi = zPc;
ca2=2*ca;
miD=zeros(ca,ca); miP=zeros(ca,ca); miDP=zeros(ca,ca);
miH=zeros(ca2,ca2); miE=zeros(ca2,ca2);
MIDa=ZEROS(CA,CA); MIPa=ZEROS(CA,CA);
u=1;
Do while u<=dimen;
miD = miD + zDi[u,].*(netaH[u]^2).*zDi[u,];
miP = miP + zPi[u,].*(netaH[u]^2).*zPi[u,];
miDP = miDP + zDi[u,].*(netaH[u].*netaH[u]).*zPi[u,];
miH = miH +
(zDi[u,]'|zPi[u,]).*(netaH[u]^2).*(zDi[u,]~zPi[u,]);
MIDA = MIDA + ZDI[U,].*(NETAD[U]^2).*ZDI[U,];
MIPA = MIPA + ZPI[U,].*(NETAP[U]^2).*ZPI[U,];
u = u + 1;
endo;

miD = miD./dimen;
MiP = miP./dimen;
MiDP = miDP./dimen;
MiH = miH./dimen;
MIDA = MIDA./DIMEN;
MIPA = MIPA./DIMEN;
sD = (AA1 * miD * AA2)./ dimen;
SDA=(AA1*MIDA*AA2)./DIMEN;
sP = (BB1 * miP * BB2) ./ dimen;
SPA=(BB1*MIPA*BB2)./DIMEN;
sDP = (AA1 * miDP * BB2) ./ dimen;
setoil = (sD~sDP)|(sDP'~sP);
wetoil = inv( cnou' * Pinv(setoil) * cnou ) * cnou' * Pinv(setoil);
betaE = wetoil * (betaD|betaP);
BETAe= BETAe +
INV(CNOU'*PINV(SETOIL)*CNOU)*(1~XBARRE)*INV((1~X
BARRE)*
INV(CNOU'*PINV(SETOIL)*CNOU)*(1~XBARRE))
*(YBARRE-(1~XBARRE)*BETAe);
netaE = y - x*betaE;

betaEm = betaEm+betaE;
betaEv = betaE.*betaE;
betaEv = betaEv+betaEv;

If ftoil > niconH0; betaPTH = betaL; betaPTE = betaL; iflag=0;
endif;

```

```

If ftoil <= niconH0; betaPTH = betaH; betaPTE = betaE; iflag=1;
endif;
if bonde==iflag; bonchoix=bonchoix+1; endif;
bPTHm = bPTHm+betaPTH;
bPTHvv = betaPTH.*betaPTH;
bPTHv = bPTHv+bPTHvv;

```

```

bPTEm = bPTEm+betaPTE;
bPTEvv = betaPTE.*betaPTE;
bPTEv = bPTEv+bPTEvv;

```

```

sigmaL = (epe)/(dimen-ca);
sigmaD = (netaD.*netaD)/(dimen-ca);
sigmaP = (netaP.*netaP)/(dimen-ca);
sigmaH = (netaH.*netaH)/(dimen-ca);
sigmaE = (netaE.*netaE)/(dimen-ca);

```

```

ssL = qL*qL'; ssLnou= qLnou*qLnou';
ssD = qD*qD'; ssDnou= qDnou*qDnou';
ssP = qP*qP'; ssPnou= qPnou*qPnou';
ssDP = qD*qP'; ssDPnou= qDnou*qPnou';
SS = (ssD~ssDP)/(ssDP~ssP); SSnou=
(ssDnou~ssDPnou)/(ssDPnou~ssPnou);

```

```
VbLh0 = diag(ssL).*sigmaL;
```

```

covbLh0=ssL.*sigmaL;
VbLh0co=xbarre.*covbLh0*xbarre' + sigmaL/dimen;
VbLh0=VbLh0co|VbLh0;

```

```
VbDh0 = diag(ssD).*sigmaD;
```

```

covbDh0=ssD.*sigmaD;
VbDh0co=xbarre.*covbDh0*xbarre' + sigmaD/dimen;
VbDh0=VbDh0co|VbDh0;

```

```
VbPh0 = diag(ssP).*sigmaP;
```

```

covbPh0=ssP.*sigmaP;
VbPh0co=xbarre.*covbPh0*xbarre' + sigmaP/dimen;
VbPh0=VbPh0co|VbPh0;

```

```
VbHh0 = diag(inv(c.*inv(ss)*c)).*sigmaH;
```

```

covbHh0=inv(c.*inv(ss)*c).*sigmaH;
VbHh0co=xbarre.*covbHh0*xbarre' + sigmaH/dimen;
VbHh0=VbHh0co|VbHh0;

```

```
VbDh1 = diag(sDA);
```

```
VbPh1 = diag(sPA);
```

```
MPINV=PINV(SSNOU);
```

```

VbHh1 =
diag(inv(cnou.*MPINV*cnou)*cnou.*MPINV*CC1.*MiH*
CC1.*MPINV*cnou*inv(cnou.*MPINV*cnou))./dimen;

```

```

MORCEAU1=INV(CNOU.*PINV(SETOIL)*CNOU);
MORCEAU2=(1~XBARRE);
MORCEAU3=INV(MORCEAU2*MORCEAU1*MORCEAU2');
MORCEAU4=INV(ZDNOU.*X/DIMEN);
MORCEAU5=INV(ZPNOU.*X/DIMEN);
MORCEAU6=(MORCEAU4~ZEROS(CA,CA))/(ZEROS(CA,CA)
~MORCEAU5);
MORCEAU7=CNOU.*PINV(SETOIL);

```

```

GRANTA=(CCNOU-
MORCEAU1*MORCEAU2'*MORCEAU3*MORCEAU2)*MOR
CEAU1*MORCEAU7

```

```

*MORCEAU6;
MORCEAU8=ZEROS(CA2,CA2);
NETAEALT=Y~X*BETAE;
MORCEAU9=1;
DO WHILE MORCEAU9<=DIMEN;
MORCO10=NETAEALT[MORCEAU9]^2;
MORCEAU8=MORCEAU8+(ZDNOU[MORCEAU9,.]|ZPNOU[
MORCEAU9,.]')*MORCO10.*
(ZDNOU[MORCEAU9,.]~ZPNOU[MORCEAU9,.]');
MORCEAU9=MORCEAU9+1;
END;
MORCEAU8=MORCEAU8./DIMEN;
VBEH1=DIAG((1/N)*(GRANTA*MORCEAU8*GRANTA));

```

```

MORCO11=(MORCEAU4*ZDNOU)/(MORCEAU5*ZPNOU');
GRANDB=((CCNOU-
MORCEAU1*MORCEAU2'*MORCEAU3*MORCEAU2)*MOR
CEAU1*MORCEAU7
*MORCO11)/(SQRT(DIMEN));
VBEH0=DIAG((1/DIMEN).*GRANDB*GRANDB'.*SIGMAE);

```

```

vbptH = vbLh0*(1-iflag)+vbHh1*iflag;
vbptE = vbLh0*(1-iflag)+vbeh1*iflag;
sVbLh0 = sVbLh0 + VbLh0; sVbDh0 = sVbDh0 + VbDh0;
sVbPh0 = sVbPh0 + VbPh0;
sVbHh0 = sVbHh0 + VbHh0; sVbDh1 = sVbDh1 + VbDh1;
sVbPh1 = sVbPh1 + VbPh1;
sVbHh1 = sVbHh1 + VbHh1; sVbEh0 = sVbEh0 + VbEh0;
sVbEh1 = sVbEh1 + VbEh1;
sVbptH = sVbptH+vbptH; sVbptE = sVbptE+vbptE;
EbLh0 = sqrt(vbLh0);
EbDh0 = sqrt(vbDh0);
EbPh0 = sqrt(vbPh0);
EbHh0 = sqrt(vbHh0);
EbEh0 = sqrt(vbEh0);
EbDh1 = sqrt(vbDh1);
EbPh1 = sqrt(vbPh1);
EbHh1 = sqrt(vbHh1);
EbEh1 = sqrt(vbEh1);
EbptH=sqrt(VbptH);EbptE=sqrt(VbptE);
sEbLh0 = sEbLh0 + EbLh0; sEbDh0 = sEbDh0 + EbDh0;
sEbPh0 = sEbPh0 + EbPh0;
sEbHh0 = sEbHh0 + EbHh0; sEbDh1 = sEbDh1 + EbDh1;
sEbPh1 = sEbPh1 + EbPh1;
sEbHh1 = sEbHh1 + EbHh1; sEbEh1 = sEbEh1 + EbEh1;
sEbEh0 = sEbEh0 + EbEh0;
sEbptH=sEbptH+EbptH; sEbptE=sEbptE+EbptE;
degre = dimen-ca;
but = nivconf/2;
bbut = ones(ca,1);
but = bbut*but;
tes = zeros(ca,1);

```

```

tbL = (betaL-vb)/(sqrt(VbLh0)); tbL=abs(tbL); tbL =
cdftc(tbL,degre);
tbD = (betaD-vb)/(sqrt(VbDh0)); tbD=abs(tbD); tbD =
cdftc(tbD,degre);
tbP = (betaP-vb)/(sqrt(VbPh0)); tbP=abs(tbP); tbP =
cdftc(tbP,degre);
tbH = (betaH-vb)/(sqrt(VbHh0)); tbH=abs(tbH); tbH =
cdftc(tbH,degre);
tbE = (betaE-vb)/(sqrt(VbEh0)); tbE=abs(tbE); tbE =
cdftc(tbE,degre);
tbptH = (betaptH-vb)/(sqrt(vbptH));tbptH=abs(tbptH);
tbptH = cdfc(tbptH,degre);

```

```

tbptE = (betaptE-vb)/(sqrt(vbptE));tbptE=abs(tbptE);
tbptE = cdfc(tbptE,degre);
tes = tBL.>but; suctbLh0 = suctbLh0 + tes;
tes = tBD.>but; suctbDh0 = suctbDh0 + tes;
tes = tBP.>but; suctbPh0 = suctbPh0 + tes;
tes = tBH.>but; suctbHh0 = suctbHh0 + tes;
tes = tBE.>but; suctbEh0 = suctbEh0 + tes;
tes=tbpth.>but;suctbpth=suctbpth+tes;
tes=tbpte.>but;suctbpte=suctbpte+tes;
tbD = (betaD-vb)/(sqrt(VbDh1)); tbD=abs(tbD); tbD =
cdfc(tbD,degre);
tbP = (betaP-vb)/(sqrt(VbPh1)); tbP=abs(tbP); tbP =
cdfc(tbP,degre);
tbH = (betaH-vb)/(sqrt(VbHh1)); tbH=abs(tbH); tbH =
cdfc(tbH,degre);
tbE = (betaE-vb)/(sqrt(VbEh1)); tbE=abs(tbE); tbE =
cdfc(tbE,degre);
tes = tBD.>but; suctbDh1 = suctbDh1 + tes;
tes = tBP.>but; suctbPh1 = suctbPh1 + tes;
tes = tBH.>but; suctbHh1 = suctbHh1 + tes;
tes = tBE.>but; suctbEh1 = suctbEh1 + tes;

" ";-----fin de la boucle numero:" r1;

r1=r1+1;

endo;

/* RESULTATS FINAUX */

format /ml/ro 11,5;

beta1=" ";

accep = acephoH;
proacepH = (acephoH/(r1-1))*100;
pourcen = proacepH;
beta1 = beta1~"BETAH";

" "; "$beta1;" ";
"NOMBRE D'ACCEPTATIONS DE H0 :"; " "; accep;" ";
"POURCENTAGE D'ACCEPTATION DE H0 :"; " "; pourcen;" ";
"puissance du test en pourcentage:" (100-proacepH);

mbL = betaLm/(r1-1);
vbetaL = (betaLv-(betaLm.*betaLm/(r1-1)))/(r1-2);
ecartL = sqrt(vbetaL);
eqmL = ((vb-mbL).*(vb-mbL))+vbetaL;
seqmL = sqrt(eqmL);
biaisL = mbL-vb;
VbLh0 = sVbLh0/nvec;
EbLh0 = sEbLh0/nvec;
typeLh0 = (1-suctbLh0/nvec)*100;

mbD = betaDm/(r1-1);
vbetaD = (betaDv-(betaDm.*betaDm/(r1-1)))/(r1-2);
ecartD = sqrt(vbetaD);
eqmD = ((vb-mbD).*(vb-mbD))+vbetaD;
seqmD = sqrt(eqmD);
biaisD = mbD-vb;
VbDh0 = sVbDh0/nvec;
EbDh0 = sEbDh0/nvec;
VbDh1 = sVbDh1/nvec;
EbDh1 = sEbDh1/nvec;
typeDh0 = (1-suctbDh0/nvec)*100;
typeDh1 = (1-suctbDh1/nvec)*100;

mbP = betaPm/(r1-1);

```

```

vbetaP = (betaPv-(betaPm.*betaPm/(r1-1)))/(r1-2);
ecartP = sqrt(vbetaP);
eqmP = ((vb-mbP).*(vb-mbP))+vbetaP;
seqmP = sqrt(eqmP);
biaisP = mbP-vb;
VbPh0 = sVbPh0/nvec;
EbPh0 = sEbPh0/nvec;
VbPh1 = sVbPh1/nvec;
EbPh1 = sEbPh1/nvec;
typePh0 = (1-suctbPh0/nvec)*100;
typePh1 = (1-suctbPh1/nvec)*100;

```

```

mbH = betaHm/(r1-1);
vbetaH = (betaHv-(betaHm.*betaHm/(r1-1)))/(r1-2);
ecartH = sqrt(vbetaH);
eqmH = ((vb-mbH).*(vb-mbH))+vbetaH;
seqmH = sqrt(eqmH);
biaisH = mbH-vb;
VbHh0 = sVbHh0/nvec;
EbHh0 = sEbHh0/nvec;
VbHh1 = sVbHh1/nvec;
EbHh1 = sEbHh1/nvec;
typeHh0 = (1-suctbHh0/nvec)*100;
typeHh1 = (1-suctbHh1/nvec)*100;

```

```

mbE = betaEm/(r1-1);
vbetaE = (betaEv-(betaEm.*betaEm/(r1-1)))/(r1-2);
ecartE = sqrt(vbetaE);
eqmE = ((vb-mbE).*(vb-mbE))+vbetaE;
seqmE = sqrt(eqmE);
biaisE = mbE-vb;
VbEh0 = sVbEh0/nvec;
EbEh0 = sEbEh0/nvec;
VbEh1 = sVbEh1/nvec;
EbEh1 = sEbEh1/nvec;
typeEh0 = (1-suctbEh0/nvec)*100;
typeEh1 = (1-suctbEh1/nvec)*100;

```

```

mbPTH = bPTHm/(r1-1);
vbPTH = (bPTHv-(bPTHm.*bPTHm/(r1-1)))/(r1-2);
ecartPTH = sqrt(vbPTH);
eqmPTH = ((vb-mbPTH).*(vb-mbPTH))+vbPTH;
seqmPTH = sqrt(eqmPTH);
biaisPTH = mbPTH-vb;
vBptH=sVbptH/nvec;
ebptH=sEbptH/nvec;
typePTH= (1-suctbptH/nvec)*100;

```

```

mbPTE = bPTEm/(r1-1);
vbPTE = (bPTEv-(bPTEm.*bPTEm/(r1-1)))/(r1-2);
ecartPTE = sqrt(vbPTE);
eqmPTE = ((vb-mbPTE).*(vb-mbPTE))+vbPTE;
seqmPTE = sqrt(eqmPTE);
biaisPTE = mbPTE-vb;
VbptE=sVbptE/nvec;
EbptE=sEbptE/nvec;
typePTE=(1-suctbptE/nvec)*100;

```

```

beta1 = "BETAL"~"BETAD"~"BETAP"~"BETAH"~"BETAE";
beta2 = "BETAPTH"~"BETAPTE";
moyenb1 = mbL~mbD~mbP~mbH~mbE;
moyenb2 = mbPTH~mbPTE;
varb1 = vbetaL~vbetaD~vbetaP~vbetaH~vbetaE;
varb2 = vbPTH~vbPTE;
ecart1 = ecartL~ecartD~ecartP~ecartH~ecartE;
ecart2 = ecartPTH~ecartPTE;
eqm1 = eqmL~eqmD~eqmP~eqmH~eqmE;
eqm2 = eqmPTH~eqmPTE;
biais1 = biaisL~biaisD~biaisP~biaisH~biaisE;
biais2 = biaisPTH~biaisPTE;

```

