



Ethical quotation system

**COVALENCE ANALYST PAPERS - CORRELATION SERIES 1**

**L'information éthique influence-t-elle le marché boursier ?  
Application au Système de cotation éthique EthicalQuote de Covalence**

David Dar Ziv, licencié en économie monétaire et financière Université de Genève  
(Suisse), stagiaire analyste, Covalence SA, Genève, 5 mai 2005

Covalence SA, Av. Industrielle 1, 1227 Carouge, Geneva, Switzerland  
tel: +41 (0)22 800 08 55 ; Fax: +41 (0)22 800 08 56  
US Representative Office, 143 Dudley Street, Cambridge MA, 02140, USA, tel +(1) 617 429 4758  
Scandinavian Rep.Office, Lokföraregatan 7C, 222 37, Lund, Sweden, Tel: 0046 (0) 46 14 97 15  
<http://www.covalence.ch> / email [info@covalence.ch](mailto:info@covalence.ch)

# **L'information éthique influence-t-elle le marché boursier ?**

**Application au  
Système de cotation éthique EthicalQuote  
de Covalence**

Projet préparé par David Dar-Ziv (stagiaire à Covalence) avec la collaboration de Matthias Brunner (Directeur de Data Doxa) et Antoine Mach (Directeur de Covalence)

Genève, Avril 2005

## **1. Introduction**

La série de faillites, entachées de nombreux scandales, de grandes entreprises américaines au cours de l'année 2002 a provoqué une véritable prise de conscience générale. La crise qui en a découlé a ramené au premier plan la question de la place à accorder aux valeurs morales dans la gestion des entreprises. A tel point qu'aujourd'hui on ne s'imagine pas une compagnie sans comportement éthique intégré comme outil de management.

Que ce soit par du sponsoring, par l'amélioration des conditions de travail de leurs employés ou par la réduction de rejets toxiques dans l'atmosphère, pour ne citer que quelques exemples, les compagnies ont compris (certaines plus tôt que d'autres) l'intérêt de telles actions : contribuer d'une manière positive à leur image publique mais aussi améliorer leur productivité ou encore faire plus de bénéfices.

Pourtant bien des pratiques non conformes à une bonne gouvernance - les activités de responsabilités sociales d'une entreprise, réparties dans les domaines économiques, sociaux et environnementaux - demeurent.

Qui fait quoi ? Qui ne fait pas ? Qui ne fait pas assez ? Comment le savoir ? Pour répondre à toutes ces questions Covalence a développé un système de cotation éthique des entreprises. Elles les suit « à la trace », compilant toutes les informations éthiques les concernant qui sont publiées dans le domaine de l'information.

Mais pour connaître l'importance de leur travail, les responsables de Covalence ont eu besoin d'estimer l'impact des informations qu'ils récoltent à travers leur système de cotation éthique. De là est né l'idée de ce projet : déterminer l'impact et le sens de l'information éthique sur le marché des grandes compagnies internationales, le marché boursier.

Avant de commencer cette recherche, Covalence et ses activités seront présentées dans le chapitre 2. Puis le chapitre 3 montrera brièvement la méthode utilisée pour réunir les données nécessaires à cette étude. Les chapitres 4 et 5 seront le corps de l'analyse statistique. Elles comporteront chacune un cas illustratif puis les résultats généraux. Et le chapitre 6 fera figure de conclusion générale.

## **2. Covalence**

### **2.1 L'entreprise**

Covalence<sup>1</sup> a été fondée en 2001 à Genève sous la forme d'une société anonyme par six personnes issues de la finance et des sciences sociales.

Covalence gère un système de cotation éthique : EthicalQuote, une base de données participative mesurant la réputation des entreprises multinationales sur les enjeux éthiques. Inspiré des cours boursiers, EthicalQuote est à la base de services offerts dans le domaine des investissements éthiques et durables, de l'analyse de réputation et de la gestion des risques.

Covalence produit des rapports détaillés sur des compagnies, secteurs ou enjeux qui offrent une vue synthétique des revendications et des actions liées à la responsabilité sociale de l'entreprise à l'échelle globale.

### **2.2 EthicalQuote et sa méthode de récolte d'information**

Sources

Le noyau de sources représente environ 200 sites web, les plus productifs sur les enjeux couverts. Certains de ces sites sont des portails ou des plates-formes de ressources, l'une d'entre-elles donnant accès aux archives de 5000 médias dans le monde. Une mailing-list parvenant d'entités

---

<sup>1</sup> Pour plus d'information sur Covalence, vous pouvez consulter son site internet à l'adresse suivante : <http://www.covalence.ch>

diverses contribue également à l'information utilisée pour constituer la base de données de Covalence.

Covalence classe les sources d'information en 9 groupes (consultant, entreprise, individu, gouvernement, organisation internationale, organisation non gouvernementale, syndicat, université) et 20 sous-groupes. Covalence ne considère pas certaines sources comme plus fiables que d'autres. Chaque source est considérée également et le réservoir est ouvert. L'essentiel est de suivre le cycle de vie des sources d'information et de gérer leur intégration dans le processus de recherche.

Covalence ne valide pas les sources d'information, ni le contenu de l'information, mais collecte, confronte et synthétise le maximum de documents pertinents issus de sources diversifiées. La politique de Covalence est de se mettre dans la situation d'un journal indépendant face à des propos, opinions et lettres de lecteurs : publier toute information dans la mesure où elle est pertinente et a un auteur identifié, sans pour autant adhérer à son contenu.

#### Critères

Une [liste de 45 critères](#) est utilisée afin de classer dans EthicalQuote les articles sélectionnés dans les sources. Les 45 critères sont divisés en 4 groupes : Conditions de travail, Impact de la production, Impact du produit, Impact institutionnel. Les références légales sont : la Déclaration universelle des droits de l'homme, les Principes directeurs de l'OCDE à l'attention des entreprises multinationales, la Déclaration de principes tripartite sur les entreprises multinationales et la politique sociale de l'OIT, la Déclaration de Rio sur l'environnement et le développement, les accords du Sommet mondial pour le développement social et le Pacte mondial des Nations Unies.

Les critères de Covalence ne sont pas différenciés selon les secteurs d'activité. Ils sont au contraire conçus pour couvrir n'importe quelle entreprise multinationale et permettre des comparaisons inter secteurs.

#### Saisie

La saisie de documents est effectuée par des stagiaires qui représentent jusqu'à présent 5 continents, 13 différents pays et de multiples filières d'études. Leur travail est précieux car il représente la base de données de Covalence, sur laquelle repose son fonctionnement.

L'intégration d'une nouvelle information dans la base de données EthicalQuote se déroule ainsi :

Sur un site sélectionné, un article traitant d'une entreprise choisie dans un des secteurs est lu par le stagiaire.

Le stagiaire décide si l'information est pertinente par rapport aux 45 critères définis. Dans l'affirmative, il sélectionne les parties du texte les plus significatives et utilise un voir deux critères au maximum pour évaluer l'article. Il attribuera ainsi à l'entreprise, à travers EthicalQuote, une note de +1 ou +2 dans le cas d'une *offre* (article positif envers l'entreprise), -1 ou -2 dans le cas d'une *demande*.

L'agrégation de toutes les entrées d'articles sont gérées à l'arrière par l'informatique. Ceci amène aux courbes d'évolution de la valeur informative des entreprises cotées par Covalence. Il faut noter que le tournus constant des stagiaires ainsi que le fait que toutes les entrées font l'objet d'une révision par une personne engagée à cet effet, permet aux données d'être « neutres » ou sans biais.

Il y a sur le site internet de Covalence l'opportunité de voir les résultats en terme de graphiques et de fiche d'entrée de documents. Si cela vous intéresse suivez le lien <http://www.covalence.ch/ethicalquote/index.php?value=0>.

### 3. Mise en place des données

On a sélectionné 10 entreprises dans 10 secteurs. Ces entreprises sont celles ayant le plus grand volume d'information dans la base de données de Covalence pendant la période juin 1999 – novembre 2004. Les secteurs économiques sont par ordre alphabétique :

Automobiles & Parts  
Banks  
Chemicals  
Entertainment & Leisure  
Food & Beverage

Mining & Metals,  
Oil & Gas  
Pharmaceuticals  
Retailers  
Technology Hardware

Voir l'[Annexe 1](#) pour la liste de toutes les entreprises.

Puis on a rassemblé la valeur éthique de chaque entreprise sélectionnée pour la période considérée depuis EthicalQuote. Et on a recherché les cours boursiers correspondants sur internet sur le site de Yahoo Finance <<http://fr.finance.yahoo.com/>>. Le processus de sélection du cours boursier s'est fait de façon systématique. La préférence a été donnée au cours local de l'entreprise (place boursière du pays où se trouve la maison mère ; par exemple pour UBS la place boursière sélectionnée est VEXTRA), puis au New York Stock Exchange

Dans le cas d'un échec de ces deux options, l'entreprise ne fait plus partie du test. Cas de Boehringer Ingelheim (non-côtée sur le marché publique, remplacée par Johnson&Johnson), Mitsui et Sumitomo.

Le cours utilisé est un cours qui a été ajusté pour neutraliser l'effet de diverses procédures menant à une variation du cours, telles que versement de dividendes ou split.

L'importation de toutes les données s'est faite à partir d'un fichier Excel directement sur SPSS (logiciel statistique). Le nombre final d'entreprises est 96. Chaque entreprise a une variable éthique (*VARIABLEe*) et boursière (*VARIABLEs*). Ce qui nous conduit à avoir 192 cours. Pour UBS par exemple, les deux variables sont codées respectivement sous la forme UBSe et UBSs.

L'analyse sera séparée en deux parties :

La première partie – *Analyse 1* - utilise la technique des moindres carrés ordinaires dans le cadre d'un modèle linéaire simple. Avec comme variable dépendante la variable boursière (VB), et la variable dépendante ou explicative sera une variable éthique (VE) , de retard le cas échéant-leading indicator. Les hypothèses liées au modèle seront brièvement explicitées et seront en parties utilisées pour tester les résultats. La deuxième partie – *Analyse 2* - prendra en compte les résultats et conclusions de l'Analyse 1 et utilisera le processus d'auto-régression sur les erreurs comme modèle d'explication.

Chaque partie sera construite de façon à suivre la procédure qui a été utilisée dans SPSS. Cela veut dire que pour chaque élément de syntaxe de SPSS il y aura une section relative dans l'analyse, permettant d'expliquer la nature et l'objectif de la syntaxe et d'explicitier ensuite les résultats.

Au niveau des fichiers SPSS : le fichier données utilisé est ***Données Corrélations COVALENCE 0.sav*** ; le fichier de syntaxe est ***Syntax Analyse 1.SPS et Syntax Scatter Plots.SPS*** ; et le fichier output est ***Output Analyse 1.spo***.

#### **4. Analyse 1 : Régression avec « leading indicator »**

##### **4.1 Les éléments de la syntaxe**

###### **Le diagramme de dispersion**

La première chose à faire lorsqu'on a des données c'est de les reporter sur un graphique. Le diagramme de dispersion est un graphique très commode pour représenter les observations simultanées de deux variables quantitatives, dans le plan générer par celles-ci. L'ensemble de ces points donne en général une idée assez bonne de la variation conjointe des deux variables et est appelé *nuage*.

Chaque symbole représente une unité d'observation et sa position dépend des valeurs obtenues pour les variables X (en abscisse) et Y (en ordonnée). Ce diagramme aide à déterminer s'il existe une relation entre les deux variables et, si oui, il aide à choisir le type d'équation qui permettrait de décrire cette relation (droite, parabole, etc.).

## La séparation des données

Il est surtout fait usage de cette technique pour analyser des séries temporelles. L'ensemble des données est réparti en deux. La période d'estimation rassemble les observations nécessaires à l'estimation des paramètres d'un modèle. La période de validation permet la vérification de la qualité d'ajustement du modèle.

Dans cette étude, la période d'estimation est Juin 1999 – Janvier 2003 (les premiers 2/3 des données). La période de validation regroupe le dernier tiers des données, soit de Février 2003 à Novembre 2004.

## La fonction de corrélation croisée (CCF)

La fonction d'autocorrélation décrit la similitude entre une série et elle-même décalée d'un intervalle  $k$ . Par analogie, la fonction de corrélation croisée permet d'estimer la ressemblance entre deux séries différentes, en décalant celles-ci l'une par rapport à l'autre. On définira des fonctions de covariances avec retard ou des corrélogrammes avec retard. En ce qui nous concerne, la CCF donnera le retard de la VE (explicative) qui a le plus grand pouvoir explicatif sur la VB.

## Le leading indicator

La CCF permet de le déterminer. Ici, il suffit de créer la variable dans SPSS. C'est la variable de retard pour laquelle la série boursière et la série éthique sont les mieux corrélées.

## La régression par les Moindres Carrés Ordinaires

Le modèle présente la variable dépendante (boursière)  $y$  comme fonction linéaire de la variable indépendante (éthique)  $x$ , avec une erreur  $e_i$ . Cette erreur explique l'écart entre les valeurs du modèle et les observations. Pour chaque observation, on a la relation suivante :

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 * x_i * \varepsilon_i \quad i = 1, \dots, n$$

Certaines hypothèses sont intrinsèques au modèle :

- H<sub>1</sub> : Linéarité entre X et Y
- H<sub>2</sub> : Cov (X,e) = 0
- H<sub>3</sub> : E(e) = 0
- H<sub>4</sub> : Var(e) = constante =  $\sigma^2$
- H<sub>5</sub> : Cov(e<sub>i</sub>,e<sub>j</sub>) = 0 ,  $i \neq j$
- H<sub>6</sub> :  $e \sim N(0, \sigma^2)$

Le but de la régression est de déterminer les coefficients de régression  $\hat{\beta}_0$  et  $\hat{\beta}_1$ , estimateurs sans biais des vrais coefficients  $b_0$  et  $b_1$ , respectivement. L'équation estimée aura la forme suivante :

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 * x_i + \hat{\varepsilon}_i \quad i = 1, \dots, n$$

et les résidus

$$\hat{\varepsilon}_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 * x_i)$$

Les estimateurs  $\hat{\beta}_0$  et  $\hat{\beta}_1$  sont estimés par la méthode des moindres carrés ordinaires, qui consiste à minimiser la somme des carrés des écarts (SCE). Le problème s'écrit de la manière suivante :

$$\text{Min SCE} = \underset{\hat{\beta}_0, \hat{\beta}_1}{\text{Min}} \sum_i^n y_i - (\hat{\beta}_0 + \hat{\beta}_1 * x_i)^2$$

C'est un simple problème de minimisation, résolu en prenant la dérivée première de la SCE par rapport à chaque paramètre. Cela nous permet d'obtenir  $\hat{\beta}_0$  et  $\hat{\beta}_1$ , les estimateurs des MCO des vrais paramètres  $\beta_0$  et  $\beta_1$  respectivement.

### Test sur les modèles

Les droites de régression calculées, elles permettent de d'estimer notre modèle en utilisant les outputs de SPSS et aussi en visualisant la qualité du modèle de prédiction dans la période de validation.

En d'autres termes, pour jauger le modèle, la façon stricte est de regarder l'output produit par SPSS mais aussi, d'une façon plus intuitive, de voir graphiquement si, pendant la période de validation, la séquence des prédictions du modèle sont proches des observations.

### Analyse des résidus et les hypothèses des MCO

Après avoir analysé l'output, on obtient une indication sur la validité de chaque modèle. Il s'agit alors de déterminer si les hypothèses liées aux MCO ont été respectées. Si elles ne le sont pas cela permet de donner une explication de la non validité du modèle et d'envisager une autre méthode d'analyse pour découvrir s'il existe une relation entre les deux groupes de variables.

Il est surtout primordial de voir le comportement des erreurs car elles sont la cause principale de l'échec de l'utilisation des MCO pour l'analyse d'une série temporelle. Pour ce faire on utilise une nouvelle fois le diagramme de dispersion et on vérifie qu'il n'y ait pas « autre chose que du hasard » dans la relation entre les erreurs, d'une part, et les prédictions et les observations des variables explicatives d'autre part.

On utilise aussi un auto-corrélogramme et un auto-corrélogramme partiel pour exclure une auto-corrélation des erreurs. C'est-à-dire pour exclure le fait que l'erreur passée influe sur l'erreur présente.

### Conclusions

Suite à toutes ces démarches on arrive à une conclusion. Soit notre modèle est valide. En effet, il y a assez de preuves (de résultats significatifs dans les outputs) pour accepter le modèle. Dans ce cas on a réussi à expliquer la série boursière par une relation linéaire avec la série éthique. Soit les résultats ne sont pas satisfaisants et on énonce les réserves liées à la méthode d'estimation des MCO pour nos séries temporelles. Et dans ce cas-ci on procède à une deuxième analyse (Analyse 2), en intégrant les remarques de non validation de l'Analyse 1.

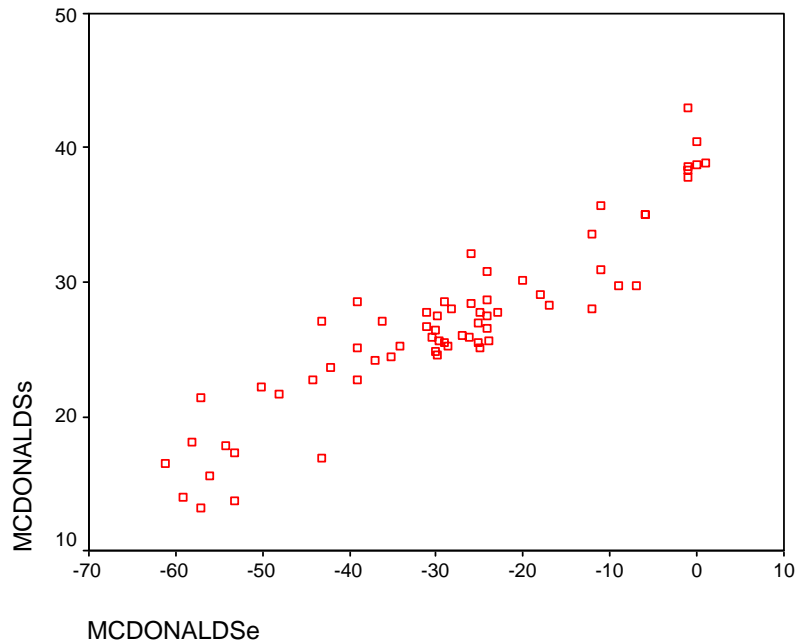
### 4.2 Les résultats des MCO

Pour simplifier la compréhension de tous les éléments liés à l'analyse ainsi que les outputs produits par SPSS, on illustrera toutes les démarches et résultats avec une des entreprises (une paire de variables). Ensuite nous discuterons de l'ensemble de l'analyse de façon plus générale.

## Cas illustratif : McDONALD'S

### Le diagramme de dispersion

Le nuage de points décrit par la représentation des observations des deux types de variables (éthique sur l'axe des abscisses, boursière sur l'axe des ordonnées) met en évidence une relation positive ( $\rho > 0$ ) et de nature linéaire (relation type  $y = ax + b$ , avec  $a > 0$ ). Les points montrent une certaine variabilité mais pas trop grande néanmoins.



De plus ils ne montrent pas de tendance à « évoluer », c'est-à-dire que la variabilité des points est constante. Donc cela nous incite à faire une régression linéaire de x (VE) sur y (VB) pour confirmer nos premières déductions. Avant cela, il nous faut déterminer le retard de la VE qui explique le mieux la VB.

### La fonction de corrélation croisée (CCF)

La CCF montre la valeur prise par les coefficients de corrélation croisée pour différents retards (-12 à +12). Dans notre cas, ce qui nous intéresse c'est la capacité de la VE à expliquer la variation de la VB. Donc seuls les « retards » positifs ou nuls permettent d'en rendre compte. Pour McDonald's le plus fort coefficient est au retard 4 ( $r = 0.191$ ). Cela veut dire que les observations de la série boursière sont mieux expliquées, au temps t, lorsque que l'on prend les observations de la série éthique au temps t-4.

Cross Correlations: MCDONALE MCDONALDSe  
MCDONALS MCDONALDSS

Transformations: difference (1)

	Cross Stand.										
Lag	Corr.	Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1
-12	-.042	.180	.	.	.	.	*	.	.	.	.
-11	-.025	.177	.	.	.	.	*	.	.	.	.
-10	.166	.174	.	.	.	.	.	***	.	.	.
-9	-.011	.171	.	.	.	.	*	.	.	.	.
-8	.079	.169	.	.	.	.	.	**	.	.	.

-7	.012	.167	.	*	.
-6	-.161	.164	.	***	.
-5	-.401	.162	**	*****	.
-4	.261	.160	.	*****	.
-3	.198	.158	.	****	.
-2	-.009	.156	.	*	.
-1	.086	.154	.	**	.
0	-.143	.152	.	***	.
1	.019	.154	.	*	.
2	-.004	.156	.	*	.
3	.148	.158	.	***	.
4	.191	.160	.	****	.
5	.004	.162	.	*	.
6	.023	.164	.	*	.
7	-.155	.167	.	***	.
8	-.015	.169	.	*	.
9	.175	.171	.	****	.
10	-.097	.174	.	**	.
11	-.025	.177	.	*	.
12	-.049	.180	.	*	.

Plot Symbols: Autocorrelations \* Two Standard Error Limits .

Toutefois, le coefficient n'est pas significatif. Explication : la CCF produit, on l'a dit, les coefficients de corrélation croisée mais aussi un seuil de signification à 95% représenté par la valeur de deux erreurs standard – la ligne en pointillé. Lorsque  $r$  ne dépasse pas cette ligne, le coefficient est dit non significatif. De manière plus formelle c'est un simple test d'hypothèse : on teste l'hypothèse nulle contre son alternative :

$$H_0 : \rho = 0$$

$$H_a : \rho \neq 0$$

Sous  $H_0$ , la probabilité que  $\rho$  soit, dans ce cas, positif est mince. Donc on ne peut pas rejeter l'hypothèse selon laquelle  $\rho$  vaut 0. A partir de là, aucun des coefficients n'est significativement différent de 0 et l'on a pas besoin d'utiliser une variable de retard. Donc afin d'expliquer la VB au temps  $t$ , on prendra la VE au temps  $t$ .

On peut maintenant procéder à la régression pour obtenir une estimation des paramètres  $\beta_1$  et  $\beta_0$  ainsi que d'autres paramètres utiles à l'analyse de la qualité du modèle.

### La régression par les Moindres Carrés Ordinaires

L'output produit par SPSS comporte plusieurs points importants. Tout d'abord le « Model Summary » nous donne des indications sur la qualité du modèle:

Model Summary(b)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.916(a)	.839	.835	2.73964003	1.029

a Predictors: (Constant), MCDONALDSe

b Dependent Variable: MCDONALDSS

Le « **R Square** » ( $R^2$ ) est un coefficient de détermination qui exprime le pourcentage de variance dans la variable dépendante, expliquée par le modèle. Un  $R^2=1$  implique que le modèle explique parfaitement (100%) la variance de la VB ; en fait le modèle prédit exactement les observations de

la VB. Un  $R^2=0$  implique que le modèle n'explique rien. Ici presque 90% (83,9% exactement) de la variation de la VB est expliquée par la variation de la VE, ce qui est plutôt bon. Pourtant on ne peut pas s'arrêter à ce seul examen.

En effet, un autre paramètre utile à regarder est la **statistique Durbin-Watson** (DW). Ce test statistique teste l'hypothèse nulle ( $H_0$ ) que les résidus des MCO ne sont pas autocorrélés contre l'hypothèse alternative qu'ils suivent un processus autorégressif d'ordre 1 ( $\varepsilon_{t-1}$  influence  $\varepsilon_t$ ). La statistique est comprise entre 0 et 4. Pour des valeurs proches de 0, les résidus sont positivement autocorrélés. Pour des valeurs proches de 4 les résidus sont négativement autocorrélés. Pour des valeurs oscillant autour de 2, il n'y a pas d'autocorrélation, bien qu'il faille le vérifier sur une table en utilisant un intervalle de valeur critiques. Ici la valeur de 1.029 indique qu'il y a une autocorrélation positive certaine des résidus. C'est un résultat déterminant car cela va à l'encontre de l'hypothèse n° 5 des MCO. On verra plus tard ce que cela implique.

Une autre statistique pour évaluer la qualité du modèle est la statistique de Fisher Snédécour, dénotée **F** dans le tableau suivant :

ANOVA(b)

Modèle		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1643.566	1	1643.566	218.978	.000(a)
	Residual	315.236	42	7.506		
	Total	1958.802	43			

Plus F est grand, en valeur absolue, meilleur est le modèle. F est pourtant plus utilisé pour comparer deux modèles qui essaient d'expliquer la variation d'une même variable dépendante.

L'examen le plus révélateur de la qualité du modèle se trouve dans le prochain tableau des coefficients:

Coefficients(a)

Modèle		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	37.507	.730		51.402	.000
	MCDONALDS e	.377	.025	.916	14.798	.000

a. Dependent Variable: MCDONALDSs

On trouve la valeur estimée des paramètres (B) du modèle avec leurs erreur standard (Std. Error) et un test de signification (t et Sig.).

**B** est la valeur prise par les estimateurs des moindres carrés  $\hat{\beta}_0$  et  $\hat{\beta}_1$ , respectivement « Constant » et « MCDONALDS e » (ce ne sont pas les vrais paramètres !). Ils indiquent quelle est la relation entre la VB et la VE. Pour une augmentation d'une unité de valeur de la VE, la VB augmenterait de 0,377.

**Std. Error** sont les erreurs standards associées aux paramètres. Elles sont utilisées afin de tester si les paramètres sont différents de 0. Elles servent aussi à calculer des intervalles de confiance des coefficients (c'est-à-dire la probabilité que la valeur vraie du coefficient soit comprise dans un intervalle défini par la valeur **B** plus ou moins (le plus souvent) 2 erreurs standard. Dans notre cas, au lieu de simplement dire que  $\hat{\beta}_1$  vaut 0.377, il peut être intéressant de nuancer le résultat. En

effet, on peut aussi affirmer, avec une probabilité de 95% que la vraie valeur du paramètre  $b_1$  est quelque part dans l'intervalle  $\hat{\beta}_1 \pm 2 * \text{Std Error}$ , donc dans l'intervalle  $0.377 \pm 0.05$ .

**t et Sig** représentent les résultats d'un test de Student qui est utilisé pour déterminer si les paramètres sont significatifs. L'hypothèse nulle ( $H_0$ : les coefficients ne sont pas significativement différents de 0) est testée contre l'hypothèse alternative ( $H_a$ : les coefficients sont significativement différents de 0).

$t$  est la valeur du  $t$  de Student associé au coefficient et  $sig.$  est la probabilité bilatérale d'obtenir une telle valeur de  $t$  si l'hypothèse nulle est vraie ( $t = B / \text{Erreur standard de } B$ ), au seuil de 95%.

La valeur de  $t$  sert à calculer si la valeur du coefficient est significativement différente de 0, c'est-à-dire si la variable explicative considérée est véritablement capable de prédire les variations de la variable expliquée.

Ici, il n'y a aucun problème à rejeter l'hypothèse nulle, c'est-à-dire à accepter que la variable MCDONALDSe prédit de façon significative les variations de la variable MCDONALDSs.

A noter : lorsque l'on a une hypothèse théorique sur le sens de la prédiction, et que le signe de  $B$  est dans la direction théoriquement prévue, on peut utiliser un test unilatéral, c'est-à-dire diviser la  $p$ -valeur par 2 avant de considérer le rejet ou non de l'hypothèse nulle.

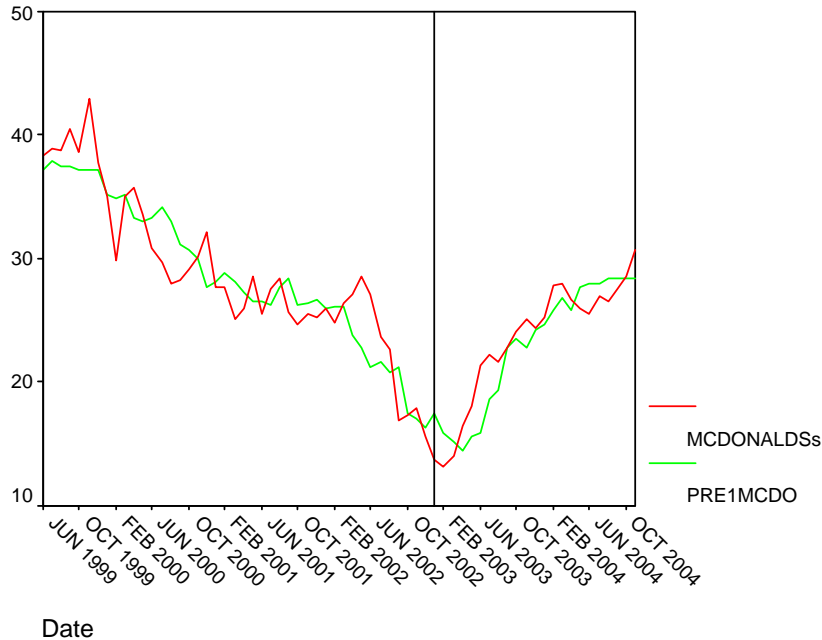
On peut donc écrire le modèle de prédiction comme ceci :

$$Y_t = 37.507 + 0.377 X_t$$

où  $y$  est la variable expliquée/boursière et  $x$  la variable explicative/éthique ; et conclure notre analyse. Mais si le modèle est idéalement bon, alors les écarts que l'on constate entre les valeurs prédites et les valeurs observées (les résidus donc) sont entièrement imputables à des erreurs de mesure. De ce fait, les résidus doivent posséder les propriétés classiques d'une distribution normale (hypothèse 6 des MCO): courbe « en cloche », symétrique autour de la valeur prédite, avec un aplatissement régulier des extrémités. En cas de violation de cette assumption, les tests de signification risquent d'être biaisés. Mais laissons cela pour la section « Analyse des résidus et hypothèses des MCO ».

Essayons de voir maintenant ce que le modèle produit comme résultats en terme de graphique. Après avoir calculé les valeurs de prédictions sur toute la période d'analyse, créant la variable du modèle « pred1mcd0 », voilà ce que cela donne :

Le graphique est séparé par une droite. A gauche de celle-ci c'est la période historique, à droite la période de validation. Dans la période historique, le modèle a un comportement globalement identique mais les variations sont moindres et décalées dans le temps. Pour la période de validation, celle qui nous intéresse en terme de validité du modèle, les variations semblent plus correspondre mais il y a toujours un décalage et dans la dernière partie du graphique le modèle s'aplatit énormément alors que la SB grimpe. On peut donc dire qu'il y a raison de croire que le modèle a des qualités explicatives mais qu'il y a quand même des choses à améliorer. Notre réponse à ces problèmes doit se trouver dans l'analyse des résidus.



### Analyse des résidus et les hypothèses des MCO

Ce sont les résidus qui vont nous révéler si nous pouvons valider notre modèle. Pour cela on doit vérifier si les hypothèses des MCO liées aux résidus tiennent la route ou si elles ne sont pas respectées.

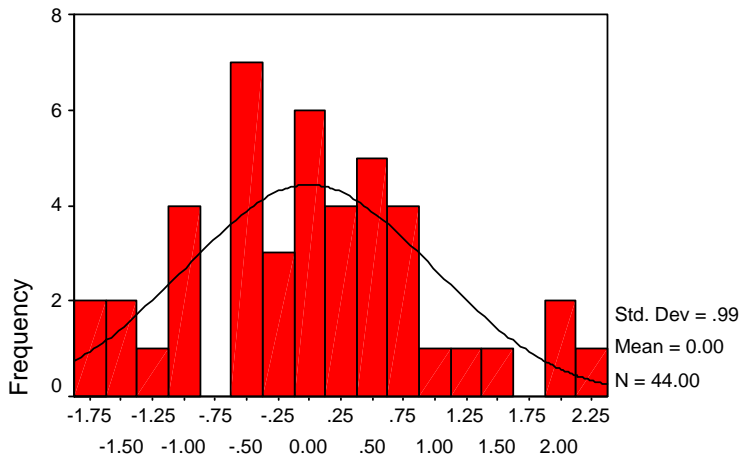
**La première approche** est de vérifier si, globalement, les résidus sont bien normalement distribués ( $H_0$  des MCO)<sup>2</sup>. Sous SPSS, on peut demander un histogramme ou un « diagramme P-P ». Le principe du premier est connu et permet immédiatement de se faire une première idée. Le principe du deuxième est de construire un diagramme mettant en relation la probabilité cumulée d'apparition d'une valeur avec sa probabilité théorique. L'histogramme P-P permet de comparer une variable avec une distribution de référence (ici on prendra évidemment la distribution normale comme référence. Idéalement, le P-P Plot ressemble à la droite identité - la diagonale).

On voit une concentration des données aux extrémités et au centre. C'est un petit écart mais cela ne remet pas en cause l'hypothèse de normalité. Voyons ce qu'il en est du P-P Plot :

<sup>2</sup> Cette étape n'est pas aussi importante qu'on le croit généralement. La normalité est une propriété qui permet aux estimateurs de converger rapidement. Le *théorème central limite* nous assure que pour des échantillons assez grands, les estimateurs que nous utilisons sont normalement distribués. La symétrie des distributions observées est un critère important qui assure une convergence rapide vers la loi normale.

## Histogram

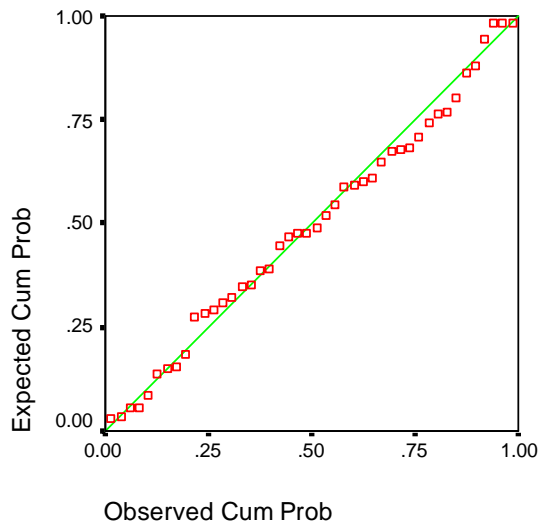
Dependent Variable: MCDONALDSs



Regression Standardized Residual

## Normal P-P Plot of Regression

Dep. Variable: MCDONALDSs



Le P-P Plot ressemble presque à la diagonale mais il y a un léger écart autour du troisième quartile (75% d'observations cumulées). Pour décider on peut aussi s'aider en utilisant un test de normalité des résidus Shapiro-Wilk. Si la valeur du Sig. est supérieure à 0.05, alors la normalité est vérifiée. C'est le cas pour les résidus de la régression sur McDonald's. On peut en conclure à la normalité des résidus.

## Tests of Normality

	Kolmogorov-Smirnov(a)	Shapiro-Wilk
--	-----------------------	--------------

	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.070	44	.200(*)	.973	44	.380

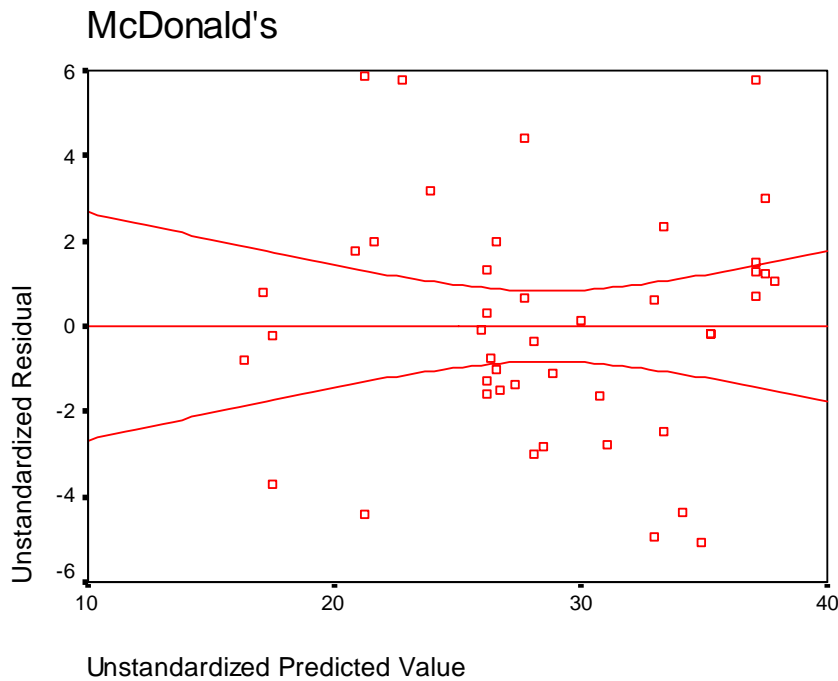
\* This is a lower bound of the true significance.

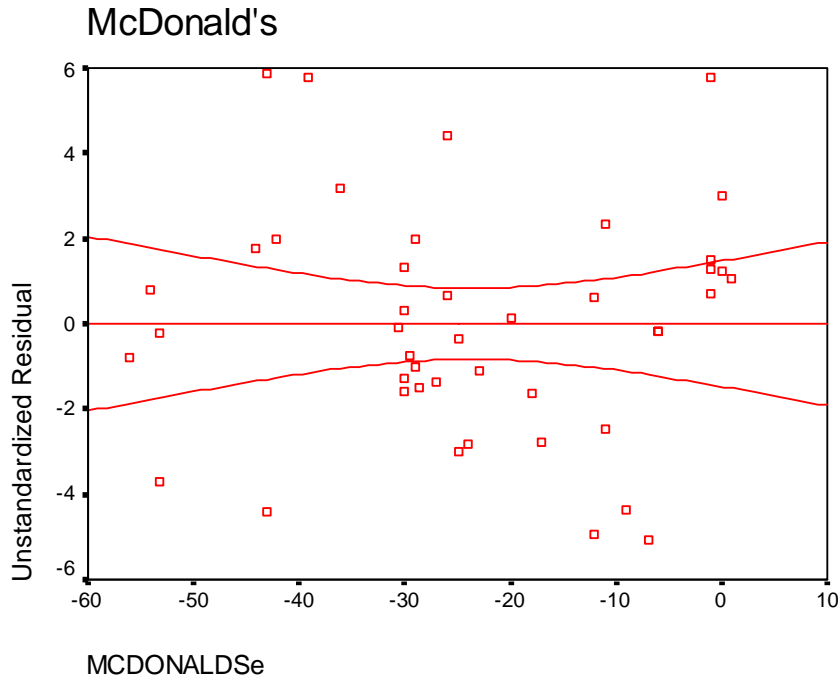
a Lilliefors Significance Correction

**La seconde approche** consiste à vérifier l'homoscédasticité des résidus - le fait que la dispersion des résidus doit être homogène sur tout le spectre des valeurs de la variable explicative d'une part, et des prédictions d'autre part ( $H_4$  des MCO). Pour cela on produit deux diagrammes de dispersion. On graphe les résidus en fonction des prédictions dans le premier, puis en fonction de la variable explicative dans le deuxième.

Si toutes les assomptions sont remplies, les résidus seront distribués presque de manière rectangulaire avec une concentration de scores le long du centre et ne montrer aucune structure repérable. On peut bien sûr dessiner des cas d'école mais sur de vraies données, comme on le voit sur nos nuages de points, tout le problème est de savoir ce que l'on entend par « presque » !

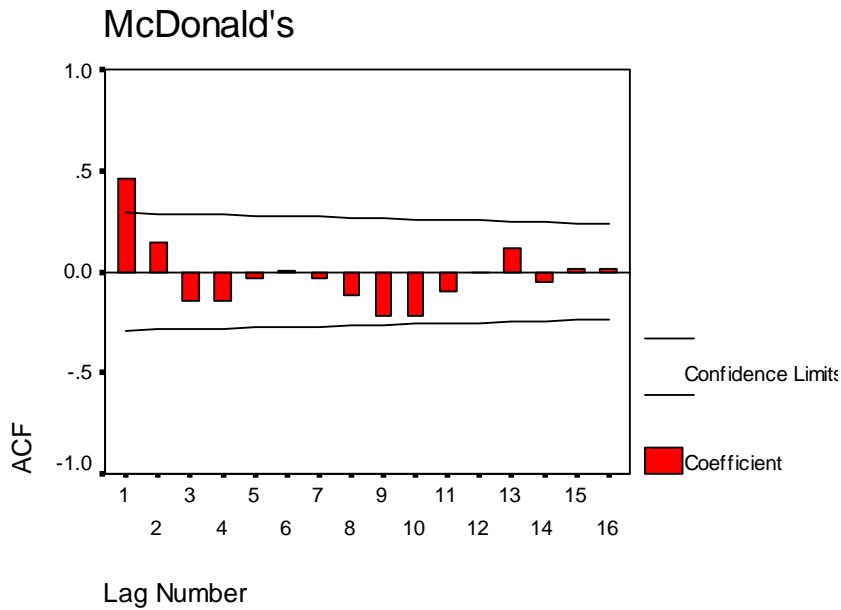
Puisque les résidus correspondent idéalement à des aléas de mesure, il n'y a pas de raison que la dispersion de ces résidus change en fonction des valeurs de la prédiction. Si la dispersion des résidus n'est pas homogène, on parle alors d'hétéroscédasticité. Sur les schémas précédents, la dispersion des résidus autour des valeurs prédites et de la variable explicative est symbolisée par les deux courbes pointillées autour de la droite. Idéalement, ces courbes devraient être parallèles à la droite. Ce n'est pas le cas, on en déduit que l'hypothèse 4 n'est pas respectée.

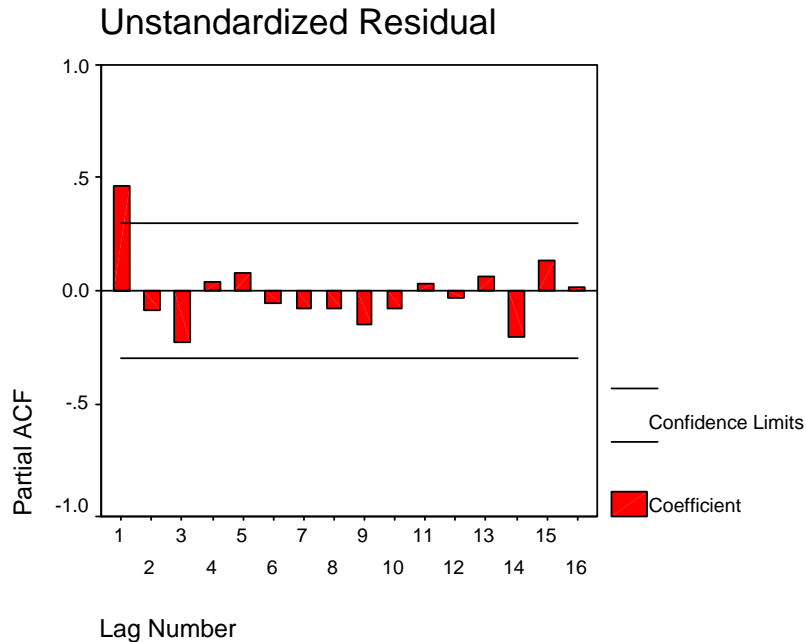




**La troisième approche** est de contrôler s'il y a une autocorrélation des erreurs. On en a déjà une indication grâce à la statistique Durbin-Watson. Sa valeur de 1.029 va en faveur d'une autocorrélation positive des résidus. Avec les fonctions d'autocorrélation (ACF) et d'autocorrélation partielle (PACF), on peut s'en assurer.

### ACF: Unstandardized Residual





L'analyse de ces fonctions est une tâche quelque peu ardue car il y a seulement des cas théoriques (parfaits), sur lesquels on se base pour déduire nos résultats. Pour approfondir ces connaissances, il est laissé le soin au lecteur de se documenter.

Nous voulons savoir si le processus des résidus est autorégressif. Cela revient à vérifier le respect ou non de l'hypothèse 5 des MCO. En simplifiant grossièrement, la littérature nous dit ceci : si la PACF n'a qu'un seul lag significatif, et que c'est le premier (les autres ne sont pas significativement différents de zéro), alors le processus est autorégressif d'ordre 1. C'est le cas ici et cela confirme le premier résultat de la statistique DW. L'hypothèse n'est donc pas respectée.

### Conclusions

Voilà l'analyse par la régression linéaire terminée. Que constate-t-on par rapport à ce que l'on pouvait attendre ?

Premièrement, on s'attendait tout d'abord à voir des relations positives dans les diagrammes de dispersion traduisant le fait que l'information éthique positive se répercute de la même façon, en terme de sens de croissance de la valeur d'une entreprise, sur le marché éthique représenté par EthicalQuote et sur le marché boursier. Or ce n'est pas le cas. En effet, 11 des 22 diagrammes montrent une relation linéaire négative. Cela se confirme dans la valeur prise par les paramètre  $\hat{b}_1$  résultants des outputs, où 12 sur 22 sont négatifs.

Deuxièmement, on pouvait croire que la force d'impact d'une information éthique, mesuré par la valeur absolue des paramètres  $\hat{\beta}_1$ , soit plus forte. Ce n'est pas le cas.

Malgré des tests sur les paramètres significatifs, un  $R^2$  fort, la statistique DW et l'analyse des résidus sont contre la validité du modèle. En effet, les hypothèses 5 et 6 de la méthode des MCO n'ont pas été respectées alors qu'elles sont un fondement pour l'utilisation de cette méthode. D'où viennent les problèmes ?

En fait, c'est la caractéristique même d'une série temporelle qui fait que la méthode des MCO n'est pas utilisable. La plupart des séries temporelles ont une tendance à monter ou descendre dans le temps, et deux séries de cette sorte seront corrélées simplement de ce fait et non pas parce qu'il existe un lien de causalité entre elles.

De plus, l'autocorrélation des erreurs (l'hypothèse d'indépendance des erreurs) est la raison principale qui ne permet pas de valider notre modèle, et qui de façon plus globale, ne permet pas l'utilisation de la méthode des MCO.

### Résultats pour toutes les séries

Ces résultats se basent sur les outputs de toutes les entreprises (voir fichier OUTPUT.doc)

### Les diagrammes de dispersion

Sur les 96 paires de séries temporelles, seules 22 nous paraissent revêtir une relation de causalité. Par secteur :

- Automobiles : **Daimler, Ford, General Motors, Nissan et Toyota**;
- Chemicals : **Air Products et Bayer AG**;
- Entertainment & Leisure : **Eastman Kodak, Matsushita et McDonald's**;
- Food & Beverage: **Starbucks**;
- Mining & Metals: **Placer Dome**;
- Oil & Gas: **ENI, Shell et Suncor Energy**;
- Pharmaceuticals : **Bayer Ag et Merck**;
- Retailers: **Carrefour et King Fisher**;
- Technology & Hardware : **IBM, Motorola et Samsung**.

Pour la plupart, les diagrammes de dispersion laissent entrevoir une relation linéaire. Celles-ci seraient :

- positives pour 8 d'entre elles : **Nissan, Air Products, Bayer, McDonald's, Starbucks, Placer Dome, Suncor Energy et Samsung**.
- Négatives pour 9 autres : **Daimler, Ford, General Motors, Eastman Kodak, Matsushita, ENI, Merck, Carrefour et IBM**.
- Autre relation pour 6 restantes : **Toyota, Air Products, Matsushita, Shell, King Fisher et Motorola**, pour lesquelles on pourrait envisager un autre type de relation (quadratique ou cubique) mais une régression linéaire sera tout de même effectuée pour celles-ci, la méthode des MCO ne s'appliquant qu'aux relations linéaires.

Une remarque : certaines entreprises sont à la fois catégorisées dans les relations linéaires et dans les autres spécifications. C'est une conséquence de leur diagramme respectif qui laisse une incertitude quant à leur spécification.

Certains diagrammes montrent peu de variance (par exemple Matsushita, Nissan, McDonald's et Starbucks) alors que d'autres y sont plus sujets (General Motors, Placer Dome ou encore IBM). On en tire pour l'instant aucune conclusion.

### La fonction de corrélation croisée (CCF)

Ici les résultats visant à créer des « leading indicators » sont pour ainsi dire inutilisables. En effet sur les 22 corrélogrammes, seuls deux ont un coefficient de corrélation significatif, permettant ainsi de créer deux leading indicators. Les 20 autres corrélogrammes induisent l'utilisation d'aucun retard dans la variable explicative.

### La régression par les Moindres Carrés Ordinaires

La méthode des MCO ne produit que des régressions significatives (sauf pour Samsung, où  $F = 3.95$ ), et seuls deux coefficients  $\hat{\beta}_1$  (liés à la variable explicative) ne sont pas significatifs. Cela veut dire que la relation de linéarité semble bien être un lien de causalité pour la majorité. Attendons quand même de voir les autres résultats.

La représentation graphique du modèle par rapport à la SB montre qu'en général les modèles ont peu de similitudes avec les SB dans la période de validation. Certains ont carrément un comportement opposé à celui de la SB (Daimler, Ford, General Motors, Toyota, Bayer Chemicals et Pharma,...). Par contre d'autres, à l'image de McDonald's ont un comportement très similaire.

### **Analyse des résidus et les hypothèses des MCO**

**Première approche.** Sur l'ensemble, l'hypothèse de normalité des résidus est trahie que trois fois (pour Merck, Motorola et Samsung). Leur P-P Plot montre clairement un écart conséquent à la diagonale et le Sig. de leur statistique Shapiro-Wilk le confirme (tous plus petit que 0.05).

**Seconde approche.** En ce qui concerne l'hypothèse d'homoscédasticité (variance constante) des résidus, les régressions sur Carrefour, Merck et Kodak semblent ne pas avoir trahi l'hypothèse. Pour celles-ci en effet il n'y a pas à vrai dire de structure repérable. Par contre pour le reste des entreprises, les diagrammes révèlent des formes d'entonnoir, caractéristique de résidus hétéroscédastiques.

**Troisième approche.** Toutes les statistiques DW sont clairement en faveur d'un processus autorégressif d'ordre 1 des erreurs. Seuls deux d'entre elles dépassent légèrement la valeur 1.

Les PACF ont toutes un lag 1 très significatif et généralement les autres lags ne le sont pas.

### **Conclusions**

Les mêmes conclusions du cas illustratif s'appliquent ici. Les paramètres sont significatifs pour la majorité des régressions. Les  $R^2$  ne sont pas utilisables car biaisés. On a beaucoup d'évidence contre la validité des modèles utilisés et en faveur d'une autorégression d'ordre 1 des résidus. On va en tenir compte dans l'analyse suivante et changer notre spécification.

## **5. Analyse 2 : Processus autorégressif sur les erreurs**

## Annexe 1

### Automobiles & Parts

BMW  
Daimler Chrysler  
Ford Motor Co.  
General Motors Corp.  
Honda Motor Co. Ltd.

Hyundai Motor Co. Ltd.  
Mitsubishi Motors Corp.  
Nissan Motor Co. Ltd.  
Toyota Motor Corp.  
Volkswagen AG

### Banks

ABN AMRO Holding N.V.  
Société Générale  
Barclays Bank  
Credit Suisse Group  
Deutsche Bank

HSBC Holdings PLC  
J.P. Morgan Chase & Co.  
Morgan Stanley  
Royal Bank of Scotland  
UBS

### Chemicals

Air Products  
Akzo Nobel  
BASF  
Bayer

Dow Chemicals  
DuPont de Nemours & Co.  
Formosa  
Mitsubishi Chemical Corp.

### Entertainment & Leisure

Eastman Kodak Co.  
LG Electronics Inc.  
Matsushita Electric  
McDonald's Corp.

Philips Electronics  
Sanyo Electric Co. Ltd.  
Sharp Corp.  
Sony Corp.

### Food & Beverage

Coca-Cola  
Danone  
Kraft Foods  
Nestlé  
PepsiCo Inc

Procter & Gamble  
SABMiller plc  
Sara Lee  
Starbucks  
Unilever

### Mining & Metals

Alcan Inc.  
Alcoa Inc.  
Anglo American  
Barrick Gold Corp.  
BHP Billiton Ltd

Gencor Ltd.  
Gold Fields Ltd.  
Newmont Mining Corp.  
Placer Dome Inc.  
Rio Tinto

### Oil & Gas

BHP Billiton Ltd  
BP PLC  
ChevronTexaco  
ENI S.P.A  
Exxon Mobil Corp

Halliburton Co  
Shell  
Statoil ASA  
Suncor Energy Inc  
TotalFinaElf

### Pharmaceuticals

Abbott  
Aventis Pasteur  
Bayer

Johnson & Johnson  
Bristol Myers Squibb  
GlaxoSmithkline

Hoffmann-La Roche  
Merck

Novartis  
Pfizer

Retailers

Boots Group PLC  
Carrefour S.A.  
Gap Inc  
H&M Hennes & Mauritz AB  
Home Depot

Kingfisher PLC  
Marks & Spencer Group PLC  
Sears, Roebuck and Co  
Target Corp.  
Wal-Mart Stores Inc.

Technology Hardware

Cisco Systems Inc.  
Dell Computer Corp.  
Ericsson  
Hewlett-Packard Co.  
IBM

Intel Corp.  
Motorola Inc.  
Nokia Corp.  
Samsung  
Siemens AG

## Annexe 2

### Liste des 45 critères (en anglais)

#### A. Working conditions

##### 1 - Labour standards

Does the company endorse the norms of the International Labour Organisation, notably those concerning freedom of association and the effective recognition of the right to collective bargaining ; the elimination of all forms of forced or compulsory labour ; the effective abolition of child labour ; and the elimination of discrimination in respect of employment and occupation (ILO Declaration on Fundamental Principles and Rights at Work)?

##### 2 - Wages

Compared to the local situation, how does the company assess the level of wages paid to employees and executives ?

##### 3 - Social benefits

Has the company taken measures internally or externally regarding social benefits and advantages for employees and families ?

##### 4 - Training and insertion

Has the company taken measures regarding youth professional training, continued formation for adults, stabilisation of jobs, social plans in case of lay-offs ? What about the cost, time and coverage for these measures ?

##### 5 - Women

Has the company taken measures regarding the promotion of women at work and the coordination of professional and private life ?

##### 6 - External working conditions

Has the company taken measures to improve working conditions among its suppliers, subcontractors and other professional partners referring to the norms of the International Labour Organization : individual initiative, code of conduct, fair trade label, social certification c.f. SA 8000?

#### B. Impact of production

##### 7 - Sales

In which countries / regions does the company sell products and services, and for what share ? (detailed information by country and type of product or brand preferred, but regional estimations also registered) What is the proportion of exports?

##### 8 - Link with official development aid

Have commercial operations of the company been related to official development assistance (bilateral or multilateral) ?

##### 9 - Export risk guarantee

Has the company benefited from export risk guarantee (ERG) ? Which where the importing countries ? Has the company answered any questionnaire relative to social and environmental aspects of export operations for authorities responsible of the ERG attribution ? Has the company published these answers ?

##### 10 - International presence

Is the impact of the company's Foreign Direct Investments positive or negative ? (FDI : Investment abroad seeking at exercising long term influence on the management of a company and representing participations from 10 % of the capital to its control and acquisition)

##### 11 - Joint ventures

Are local investors participating in the investments realized by the company, at what share ?  
Are these partnerships due to deliberate choices or to a legal constraint ?

#### **12 - Economic impact**

How does the company's investments influence local industries in terms of job creation, access to markets, competition, economic growth ?

#### **13 - Social impact**

How does the company's investments influence the implementation of local laws relating to social areas c.f. social protection, public health, employee relations, fiscal relations ?

#### **14 - Job stability**

In the different countries / regions where it is active, what is the turn-over of the company's employees ?

#### **15 - Local employees**

In the different countries / regions where it is active, what is the number and the proportion of local employees in the company ?

#### **16 - Local executives**

In the different countries / regions where it is active, what is the number and the proportion of local executives in the company ?

#### **17 - Women employed**

In the different countries / regions where it is active, what is the proportion of women among the company's employees and among the company's executives?

#### **18 - Downsizing**

Has the company sensibly reduced or interrupted activities for economic reasons ? If yes, what can be said about the measures taken to minimize negative social effects of such decisions ?

#### **19 - Infrastructures**

Have FDI realized by the company provoked governmental investments in infrastructures for transport, energy, health, education, training and other fields of public interest ? Has the company supported or directly participated to investments in public infrastructures ?

#### **20 - Local sourcing**

Does the company purchase directly to local farmers, industries and services providers ? If yes, how does the company assess the impact of these buyings on economic and social development ? How many local suppliers in each country ?

#### **21 - Stability of prices**

How does the company 's buying of raw material influence price fluctuations in international markets ?

#### **22 - Technical assistance**

Do the company's suppliers, subcontractors et other professional partners benefit from its technical assistance, advice and training ? If yes, how does the company assess the impact of such transfers of knowledge and technology on local economic and social development ?

#### **23 - Intellectual propriety rights**

While managing its intellectual propriety rights, has the company taken measures that favorize human and economic development, the protection of biodiversity, respect of traditional knowledge and local natural resources, for example through voluntary licenses, agreements, cooperation with research institutes and local communities ?

#### **24 - Local innovation**

In which measure does the company's presence influe on local industrial and cultural innovation ? Has the company's presence contributed to local innovations useful to economic and social development ?

#### **25 - Fiscal contributions**

For each country / region where it is active, what are the following datas : fiscal contributions, purchasing of raw material and products, sales, profit ? What can the company say about its fiscal relations policy ? How can the company assess the impact of its fiscal contributions to

local economic and social development ?

**26 - Environmental impact of production**

Has the company adopted programs of management of the environmental impact of its activities ? What can be said about the effect of these programmes on local economic and social development?

**C. Impact of product**

**27 - Product human risk**

Has the company had to face requests regarding risks that its products would present for the health, safety and security of consumers and populations ? If yes, how did it react to these requests ?

**28 - Product social utility**

Does the company currently market products or services that respond particularly to the needs related to human, social and economic development ?

**29 - Product relation to culture**

Has the company put on the market a product, or developed an activity, that value local culture and traditions ?

**30 - Socially innovative product**

Does the company research & develop products or services that present a particular interest for local economic and social development (R&D) ?

**31 - Product environmental risk**

Has the company taken particular measures relatively to environmental risks of certain products, c.f. reference to international agreements, cooperation with international agencies, NGOs, universities, local communities ?

**32 - Waste management**

Has the company taken particular measures relatively to the management of waste due to its products ?

**33 - Eco-innovative product**

Has the company launched a new product or service environmentally friendly while contributing to human development ?

**34 - Information to consumer**

Has the company taken particular measures aiming at informing consumers on the use of its products, c.f. price, quality, quantity needed and other human, economic, cultural or social implications ?

**35 - Pricing / needs**

Has the company taken particular measures to define moderate selling prices for products responding to essential human needs as defined by major development organizations ?

**36 - Cause related marketing**

Has the company been involved in operations of cause related marketing (a % of the sale of a product is attributed to social projects) ?

**37 - Social sponsorship**

Does the company proceed to donations or social sponsorship in cash or in kind ? In that case, does the company communicate identity of beneficiaries and the amount involved ? What is its policy in this field ? What does the company expect of NGOs to engage in more cooperation ?

**D. Institutional impact**

**38 - Anti-corruption policy**

Is the company taking measures to contribute to reduce corruption in private and public markets ?

### **39 - Humanitarian policy**

Has the company defined a policy and taken particular measures regarding its activities in conflict or reconstruction post-conflict zones, a humanitarian policy ? Has the company been requested to leave certain countries because of political situations, or through pressure on governmental or rebel forces, and in that case what was the reaction to these requests ?

### **40 - Human Rights policy**

Has the company defined a human rights policy and taken particular measures regarding its activities in the developing world ?

### **41 - United Nations policy**

Does the company support the UN Global compact or other UN initiatives, programs and agencies ? In this case, is this support direct or via a professional organisation ? Has the company taken initiatives related to the Global compact ?

### **42 - Boycott policy**

Has the company been the object of demands of exit of certain countries because of the human rights situation ("repressive regimes"), and in this case how did it respond to these demands ?

### **43 - Social stability**

Has the company taken particular measures that contribute directly to local social stability in rural and urban areas, notably in matters of education, health, security ?

### **44 - Support to political actors**

Does the company support political actors financially, is this information available to the public (parties, countries) ?

### **45 - Lobbying practices**

Which national and international political issues has the company discussed about with developing countries governments ? What are the main subjects of debate currently ?

## **Bibliographie**

SPSS Inc., SPSS Trends 10.0, 1999.

Hamilton, Time Series Analysis, 1984, MIT Press.

SPSS Inc., SPSS for Windows: Base System User's Guide, Release 6.0, 1993.