Université de Montréal

Les tests de causalité en variance entre deux séries chronologiques multivariées

par

Herbert Tchahou Nkwimi

Département de mathématiques et statistique

Université de Montréal

Faculté des arts et des sciences

Mémoire présenté à la faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences(M.Sc)
en Statistique

Décembre, 2010

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé

Les tests de causalité en variance entre deux séries chronologiques multivariées

présenté par

Herbert Tchahou Nkwimi

a été évalué par un jury composé des personnes suivantes :

Jean-François Angers
(président-rapporteur)

Pierre Duchesne
(directeur de recherche)

Roch Roy
(membre du jury)

Mémoire accepté le :

SOMMAIRE

Les modèles de séries chronologiques avec variances conditionnellement hétéroscédastiques sont devenus quasi incontournables afin de modéliser les séries chronologiques dans le contexte des données financières. Dans beaucoup d'applications, vérifier l'existence d'une relation entre deux séries chronologiques représente un enjeu important. Dans ce mémoire, nous généralisons dans plusieurs directions et dans un cadre multivarié, la procédure dévéloppée par Cheung et Ng (1996) conçue pour examiner la causalité en variance dans le cas de deux séries univariées. Reposant sur le travail de El Himdi et Roy (1997) et Duchesne (2004), nous proposons un test basé sur les matrices de corrélation croisée des résidus standardisés carrés et des produits croisés de ces résidus. Sous l'hypothèse nulle de l'absence de causalité en variance, nous établissons que les statistiques de test convergent en distribution vers des variables aléatoires khi-carrées. Dans une deuxième approche, nous définissons comme dans Ling et Li (1997) une transformation des résidus pour chaque série résiduelle vectorielle. Les statistiques de test sont construites à partir des corrélations croisées de ces résidus transformés. Dans les deux approches, des statistiques de test pour les délais individuels sont proposées ainsi que des tests de type portemanteau. Cette méthodologie est également utilisée pour déterminer la direction de la causalité en variance. Les résultats de simulation montrent que les tests proposés offrent des propriétés empiriques satisfaisantes. Une application avec des données réelles est également présentée afin d'illustrer les méthodes.
Mots clés : Causalité en variance ; variances conditionnellement hétérocédastiques ; tests portemanteaux ; corrélations croisées résiduelles ; séries chronologiques multivariées.

SUMMARY

Time series models with conditionnaly heteroskedastic variances have become almost inevitable to model financial time series. In many applications, to confirm the existence of a relationship between two time series is very important. In this Master thesis, we generalize in several directions and in a multivariate framework, the method developed by Cheung and Ng (1996) designed to examine causality in variance in the case of two univariate series. Based on the work of El Himdi and Roy (1997) and Duchesne (2004), we propose a test based on residual cross-correlation matrices of squared residuals and cross-products of these residuals. Under the null hypothesis of no causality in variance, we establish that the test statistics converge in distribution to chi-square random variables. In a second approach, we define as in Ling and Li (1997) a transformation of the residuals for each residual time series. The test statistics are built from the cross-correlations of these transformed residuals. In both approaches, test statistics at individual lags are presented and also portmanteau-type test statistics. That methodology is also used to determine the direction of causality in variance. The simulation results show that the proposed tests provide satisfactory empirical properties. An application with real data is also presented to illustrate the methods.
Key words : Causality in variance ; Conditional heteroscedasticity ; Portmanteau test statistics ; Residual cross-correlations ; Multivariate time series.

Remerciements

À ma grand mère

Je voudrai dans un premier temps remercier Pierre Duchesne qui a bien voulu diriger ce travail de recherche. Ses conseils judicieux, sa rigueur, sa disponibilité ainsi que son aide financière ont été déterminant pour l'aboutissement de ce travail.

Je voudrai dédier ce travail à ma grand mère qui nous a quitté quelques mois plus tôt ; *"maman Elisabeth tu resteras à jamais gravé dans nos coeurs ; repose en paix"*.

Je tiens aussi à remercier ma famille, mon épouse Bertille Tchanda pour sa patience, mon père Jean Tchahou pour tout son soutien depuis de longues années, ma mère Désirée Nyandja pour ses encouragements ainsi que tous mes frères et soeurs.

Que tous ceux qui de près ou de loin ont contribués à l'aboutissement de ce travail reçoivent ma reconnaissance et ma gratitude.

Table des matières

INTRODUCTION **1**

1

Préliminaires **5**

1.1 Séries temporelles 5

1.2 Processus stationnaires 6

1.2.1 Estimation de la moyenne et des autocovariances 7

1.2.2 Modélisation des séries stationnaires 8

1.2.3 Notion de causalité et d'inversibilité d'un processus VARMA 8

1.3 Processus conditionnellement hétéroscédastiques 10

1.3.1 Modèles DVEC 11

1.3.2 Modèles BEKK 12

1.4 Notion de causalité entre deux séries chronologiques 12

1.4.1 Causalité en moyenne 12

1.4.2 Causalité en variance 14

2

Article **16**

3

Démonstration des principaux résultats **56**

3.1 Preuve de la proposition 1 56

3.2 Preuve de la proposition 2 62

4

Simulations complémentaires 67
4.1 Description des expériences 67
4.2 Discussion des résultats 68

5

Étude de cas supplémentaire 73
5.1 Description des indices et notations 73
5.2 Test de causalité en variance 75

Conclusion 82

Bibliographie 83

Table des figures

5.1 Indice du marché Suisse, indice CAC 40, indice N225 et indice HIS de janvier 2007 à décembre 2009. 74
5.2 Séries des rendements. 74

Liste des tableaux

4.1 Puissance empirique, modèle S_1, niveaux nominaux 5% et 10%, $n = 1000, 2000$, pour les tests globaux $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ et $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$. 69

4.2 Puissance empirique, modèle S_1, niveaux nominaux 5% et 10%, $n = 1000, 2000$, pour les tests aux délais individuels $\mathrm{ER}_n(k)$, $\mathrm{ER}^*_n(k)$, $\mathrm{LL}_n(k)$ et $\mathrm{LL}^*_n(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$. 70

4.3 Puissance empirique, modèle S_2, niveaux nominaux 5% et 10%, $n = 1000, 2000$, pour les tests globaux $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ et $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$. 71

4.4 Puissance empirique, modèle S_2, niveaux nominaux 5% and 10%, $n = 1000, 2000$, pour les tests aux délais individuels $\mathrm{ER}_n(k)$, $\mathrm{ER}^*_n(k)$, $\mathrm{LL}_n(k)$ et $\mathrm{LL}^*_n(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$. 72

5.1 Estimateurs pour les modèles des indices asiatiques et européens avec les valeurs t correspondantes. 77

5.2 Tests statistiques aux délais individuels appliqués sur les résidus standardisés. 77

5.3 Estimateurs des paramètres pour le modèle augmenté de l'indice asiatique avec les valeurs t correspondantes. 78

5.4 Test statistique de El Himdi et Roy (1997) pour tester la causalité en moyenne. Les valeurs P approximatives sont également données. 79

5.5 Tests statistiques globaux pour tester la causalité en variance $\mathrm{ER}^*_{n,M}$ et $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$. Les degrés de liberté et les valeurs P sont également fourni. . . 80

5.6 Tests statistiques aux délais individuels pour détecter la causalité en variance $\mathrm{ER}^*_n(k)$ et $\mathrm{LL}^*_n(k)$, $|k| \leq 10$. 81

INTRODUCTION

L'environnement économétrique s'est profondément métamorphosé au cours des deux dernières décennies. On a observé un développement des outils statistiques de plus en plus sophistiqués dans la compréhension et la gestion des marchés boursiers. Une question naturelle et d'intérêt porte sur l'étude de chocs financiers et leur propagation d'un marché boursier à un autre. En effet, il arrive fréquemment que l'effondrement (ou la hausse) d'actifs d'un marché boursier se répercute sur les rendements d'un marché voisin. Il apparaît par exemple que les marchés américains et européens sont liés. Dès lors, toute information recueillie sur l'un des actifs peut s'avérer importante dans la caractérisation de la dynamique de l'autre. On parle alors de lien de causalité entre ces deux marchés. En finance, Ross (1989) a montré qu'en l'absence de toute opportunité d'arbitrage, la variance des rendements est directement liée au flot d'information, suggérant encore que la nature d'un outil financier peut être influencée par les conditions se trouvant dans un marché. Granger (1969) fut l'un des premiers auteurs à formaliser la notion de causalité entre deux séries, communément connue sous l'appellation «lien de causalité au sens de Granger». Par exemple, considérons l'une des premières méthodes qui a permis d'étudier ce genre de problème. Afin de déterminer si une variable $\mathbf{X}_1$ cause une variable $\mathbf{X}_2$, on commence par modéliser la valeur actuelle de $\mathbf{X}_2$ au moyen d'une régression linéaire sur ses valeurs passées. Par la suite, on regarde si l'estimation s'améliore grâce à la prise en compte de la valeur actuelle de $\mathbf{X}_1$ et éventuellement de ses valeurs passées. Plus précisément, on regarde si la variable $\mathbf{X}_1$ est déterminante dans l'estimation de $\mathbf{X}_2$ en testant si les coefficients des valeurs retardées de $\mathbf{X}_1$ sont significativement différents de zéro.

De façon plus formelle, considérons le processus multivarié suivant :

$$\mathbf{X} = \{\mathbf{X}_t = (\mathbf{X}_{1t}, \mathbf{X}_{2t}) : t \in \mathbb{Z}\} = \{\mathbf{X}_1, \mathbf{X}_2\}.$$

On présume que $\mathbf{X}$ admet des moments d'ordre deux et posons $\overline{\mathbf{X}_t}$ l'espace de Hilbert engendré par les composantes de $\{\mathbf{X}_s,\ s \leq t\}$. Notons $\mathbf{P}\left(\mathbf{X}_{it}/\overline{\mathbf{X}_{t-1}}\right)$ la projection orthogonale de $\mathbf{X}_{it}$ sur $\overline{\mathbf{X}_{t-1}}$, $i = 1, 2$. Considerons la projection $\mathbf{P}\left(\mathbf{X}_{2t}/\overline{\mathbf{X}_{t-1}}\right)$ et supposons que le processus $\mathbf{X}_2$ admet une représentation autorégressive. On peut alors écrire :

$$\mathbf{X}_{2t} = \boldsymbol{\alpha}_0 + \sum_{i=1}^{\infty} \boldsymbol{\alpha}_i \mathbf{X}_{2,t-i} + \sum_{j=0}^{\infty} \boldsymbol{\beta}_j \mathbf{X}_{1,t-j} + \boldsymbol{\epsilon}_t.$$

Afin de vérifier que $\mathbf{X}_1$ ne cause pas $\mathbf{X}_2$, il semble naturel de chercher à vérifier l'hypothèse suivante :

$$\mathbf{H}_0 : \boldsymbol{\beta}_i = \mathbf{0}, \quad i = 0, 1, 2, \ldots$$

On note que dans l'hypothèse précédente, un nombre infini de coefficients $\boldsymbol{\beta}_i$ sont présents. Ainsi, en pratique, on considère des modèles tronqués :

$$\mathbf{X}_{2t} = \boldsymbol{\alpha}_0 + \sum_{i=1}^{p} \boldsymbol{\alpha}_i \mathbf{X}_{2,t-i} + \sum_{j=0}^{p} \boldsymbol{\beta}_j \mathbf{X}_{1,t-j} + \boldsymbol{\epsilon}_t,$$

où p est un entier spécifié par l'analyste. On vérifie alors :

$$\mathbf{H}_0 : \boldsymbol{\beta}_i = \mathbf{0}, \quad i = 0, 1, 2, \ldots, p.$$

Le choix de p est une considération importante. Sous des hypothèses générales, les coefficients $\boldsymbol{\alpha}_i$ devraient décroître avec i. Il faudra que p soit suffisamment grand afin de retenir les coefficients significatifs. Hamilton (1994, p.305) discute de l'importance du choix de p et son impact sur les tests de causalité. Afin de mettre en oeuvre le test, des statistiques issues de la théorie des modèles linéaires permettent d'apprécier si les coefficients $\boldsymbol{\beta}_i$ semblent significatifs. Voir par exemple Hamilton (1994, p. 304). Notons que Hamilton (1994) discute du cas où les deux séries $\{\mathbf{X}_{1t}\}$ et $\{\mathbf{X}_{2t}\}$ sont univariées. Notons également que compte tenu de la dépendance inhérante au fait que $\{\mathbf{X}_{1t}\}$ et $\{\mathbf{X}_{2t}\}$ sont des processus stochastiques, une théorie exacte n'est habituellement pas possible. Cependant, Hamilton (1994) énonce que sous des conditions de régularité, les statistiques-F pour tester l'hypothèse $\mathbf{H}_0$ précédente sont valides asymptotiquement.

Depuis le travail fondateur de Granger, plusieurs auteurs se sont penchés sur la causalité. Sims (1972) a également étudié cette question lorsque deux séries chronologiques univariées sont stationnaires au second ordre. Dans la formulation de Sims (1972), $\{\mathbf{X}_{2t}\}$ s'exprime en fonction du passé, du présent et du futur de $\{\mathbf{X}_{1t}\}$. Après une transformation correspondant essentiellement à un filtrage de la série, il est également possible de faire des tests avec des statistiques-F. Voir Hamilton (1994, p. 305).

La méthodologie utilisée dans les tests précédents repose essentiellement sur l'approche de Granger et l'estimation des coefficients d'une fonction affine. Les auteurs dans la spécification du modèle sous-entendent l'existence d'une relation linéaire entre les deux séries. Cependant dans de nombreux phénomènes, nous ne pouvons pas exclure *a priori* l'existence d'une dynamique non linéaire entre les séries. Pour cette raison, ces tests ne couvrent malheureusement pas la totalité des cas rencontrés en pratique. Afin de remédier à cette insuffisance, Hiemstra et Jones (1994) ont proposé une version du test de causalité permettant de prendre en compte les dynamiques non linéaires. Une autre préoccupation non moins importante du test de causalité selon Granger réside dans sa difficulté à déterminer la direction de la causalité. Ceci a d'ailleurs fait l'objet d'un article par Jacobs, Leamer et Ward (1979).

De nombreux articles plus récents ont fait l'objet de l'étude de la causalité en variance. Afin de tester la causalité en variance entre deux séries, Cheung et Ng (1996) ont développé une procédure de test en deux étapes. Dans une première étape, ils proposent un modèle pour chacune des séries. Les résidus standardisés sont ensuite utilisés dans une deuxième étape pour construire une statistique de test basée sur la corrélation croisée des résidus élevés au carré. Hong (2001) a généralisé les tests de Cheung et Ng (1996) en utilisant la somme pondérée des carrés des corrélations croisées entre deux séries de résidus standardisés. Il a illustré par des simulations Monte Carlo que les tests avec une pondération non uniforme sont souvent plus performants en terme de puissance que les tests de Cheung et Ng (1996) qui sont des cas particuliers en utilisant une pondération uniforme tronquée. Cependant, les niveaux empiriques de Hong (2001) sont souvent loin des niveaux nominaux. Ceci survient car les statistiques de test qui utilisent une approche spectrale convergent plutôt lentement vers leur distribution asymptotique. Notons que des méthodes de ré-échantillonnage (bootstrap paramétrique, par exemple), permettent de corriger ce genre de problèmes. Voir Poulin et Duchesne (2008).

Tous les articles ci-dessus ont abordé la notion de causalité dans un cadre univarié. Ce cadre est quelque peu restrictif, puisqu'en pratique plusieurs actifs doivent être analysés. En effet, en matière d'investissement, les actifs sont souvent regroupés au sein d'un ou plusieurs portefeuilles, qui peuvent à leur tour comprendre un ou plusieurs actifs. Dans ce dernier cas, les tests développés dans le cadre univarié deviennent alors insuffisants si on considère le portefeuille dans sa globalité et si on s'intéresse à la causalité entre deux groupes d'indices. Une gamme d'outils a donc vu le jour constituant une extension de ces notions dans un cadre multivarié.

Le présent travail s'inscrit dans cette direction. Nous nous intéresserons notamment à la causalité en variance entre deux séries chronologiques multivariées. Deux approches seront considérées dans ce travail. La première est une généralisation de l'approche de Cheung et Ng (1996) en nous appuyant sur le travail de El Himdi et Roy (1997) et utilisant la mesure de dépendance matricielle introduite dans Duchesne (2004) des résidus standardisés carrés et des produits croisés de ces résidus. Utilisant les covariances croisées résiduelles, nous construirons deux types de statistiques, une statistique de test de type portemanteau pour la causalité en variance ainsi qu'une statistique de test pour les délais individuels. Dans le cas de l'existence de causalité, ce deuxième test sera très utile dans la caractérisation de la direction de la causalité.

Dans une seconde approche, nous utiliserons la transformation proposée dans Ling et Li (1997), pour ramener les résidus vectoriels en quantités scalaires. Le test est alors basé sur les covariances croisées de ces résidus transformés. Comme dans l'approche précédente, nous construisons les tests de type portemanteau ainsi qu'un test pour les délais individuels.

La suite de ce document s'articule autour de cinq chapitres. Après un bref rappel de quelques notions importantes, le chapitre 1 présente sommairement les notions de causalité en moyenne, de causalité en variance, et ce dans un cadre multivarié. Le chapitre 2 comprend intégralement un article qui contient les résultats de notre recherche. Dans le chapitre 3, nous présentons les preuves des résultats principaux de l'article. L'article présente des résultats de simulation et une analyse de données. Cependant, les résultats contenu, dans l'article ne représentent qu'une fraction de nos expériences empiriques. Ainsi, par la suite, nous présenterons quelques résultats de simulation Monte Carlo complémentaires dans le chapitre 4. De plus, une analyse de données supplémentaire traitant cette fois-ci de l'étude de la causalité en variance entre un groupe d'indices boursiers européens et un groupe d'indices asiatiques concluera ce travail.

Chapitre 1

Préliminaires

1.1 Séries temporelles

Définition 1.
Soit $(\Omega, \mathcal{F}, P)$ *un espace probabilisé, sur lequel nous considérons une suite de variables aléatoires réelles* $\{\mathbf{X}_t,\ t \in \mathbb{Z}\}$*. Pour tout* $t \in \mathbb{Z}$*, le vecteur aléatoire* $\mathbf{X}_t$ *sur* Ω *et pour tout* $\omega \in \Omega$*,* $\mathbf{X}_t(\omega)$ *est une réalisation de* $\mathbf{X}_t$*.*

On constate dans la définition 1 que $\{\mathbf{X}_t,\ t \in \mathbb{Z}\}$ est un processus stochastique à temps discrets. Ainsi une série temporelle est une suite de variables aléatoires observées dans le temps. On parle aussi de série chronologique. La compréhension de certains phénomènes aléatoires nécessite une modélisation simultanée de plusieurs séries. Les séries chronologiques multivariées se trouvent ainsi très utiles dans de nombreuses applications. C'est d'ailleurs de ces séries dont il sera question tout au long de ce travail. Introduisons la définition suivante :

Définition 2.
Une série chronologique multivariée est définie comme une suite d'observations d'un processus stochastique multivarié $\{\mathbf{X}_t,\ t \in \mathbb{Z}\}$*. Pour* t *fixé,* $\mathbf{X}_t = (X_t(1), \ldots, X_t(d))^\top$ *représente un vecteur de dimension* d*. Sous la condition* $E\{\mathbf{X}_t^2(i)\} < \infty$*, pour tout* $t \in \mathbb{Z}$ *et tout* $i = 1, \ldots, d$*, on définit les deux premiers moments du processus par :*

$$E(\mathbf{X}_t) = \boldsymbol{\mu}_t = (\mu_t(1), \ldots, \mu_t(d))^\top, \tag{1.1}$$
$$\mathbf{\Gamma}_{\mathbf{X}}(t+h, t) = E\left\{\left(\mathbf{X}_{t+h} - \boldsymbol{\mu}_{t+h}\right)\left(\mathbf{X}_t - \boldsymbol{\mu}_t\right)^\top\right\} = \left(\gamma_{i,j}(t+h, t)\right)_{i,j=1}^d, \tag{1.2}$$

qui représentent respectivement le vecteur des moyennes et la matrice des variances-covariances entre $\mathbf{X}_{t+h}$ *et* $\mathbf{X}_t$*.*

1.2 Processus stationnaires

L'un des concepts jouant un rôle important dans l'étude des séries chronologiques est la notion de stationnarité. Ce concept permet des simplifications importantes dans l'analyse de données dépendantes et facilite définitivement l'inférence statistique. On distingue deux types de stationnarité, la stationnarité stricte et la stationnarité au sens large.

Définition 3.
Un processus $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ *de dimension* d *est stationnaire au sens strict ou strictement stationnaire si pour tout* n *entier naturel et* $t_1, \ldots, t_n$ *et* h *entiers relatifs, la distribution de* $X_{t_1}, \ldots, X_{t_n}$ *est la même que celle de* $X_{t_1+h}, \ldots, X_{t_n+h}$*.*

C'est une notion assez difficile à vérifier en pratique ; la notion suivante est beaucoup plus facile à vérifier empiriquement.

Définition 4.
Un processus discret $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ *tel que* $E\left\{\mathbf{X}_t^2(i)\right\} < \infty$*, pour tout* $t \in \mathbb{Z}$ *et tout* $i = 1, \ldots, d$ *est stationnaire au sens large ou faiblement stationnaire si les conditions suivantes sont vérifiées :*

$$E(\mathbf{X}_t) = \boldsymbol{\mu}_t \equiv \boldsymbol{\mu}, \tag{1.3}$$

$$\boldsymbol{\Gamma}_{\mathbf{X}}(t+h, t) = E\left\{(\mathbf{X}_{t+h} - \boldsymbol{\mu})(\mathbf{X}_t - \boldsymbol{\mu})^\top\right\} \equiv \boldsymbol{\Gamma}_{\mathbf{X}}(h). \tag{1.4}$$

Ainsi, la relation (1.3) est indépendante du temps et (1.4) devient indépendante de t et fonction uniquement du délai h. L'exemple le plus simple de processus stationnaire au sens large (on dit aussi au second-ordre) est celui du bruit blanc. Il est très utile dans la construction de modèles de séries chronologiques.

Définition 5.
Un processus $\{\boldsymbol{\epsilon}_t, t \in \mathbb{Z}\}$ *de dimension* d *est un bruit blanc, s'il est faiblement stationnaire de moyenne nulle et de matrice des variances-covariances :*

$$\boldsymbol{\Gamma}_{\boldsymbol{\epsilon}}(h) = \begin{cases} \boldsymbol{\Sigma}, & \text{si } h\text{=0,} \\ \mathbf{0}, & \text{sinon.} \end{cases}$$

On note souvent $\{\boldsymbol{\epsilon}_t, t \in \mathbb{Z}\} \sim BB(\mathbf{0}, \boldsymbol{\Sigma})$. Si les vecteurs $\boldsymbol{\epsilon}_t$ sont indépendants et identiquement distribués, on parle de bruit blanc fort. Il existe également une classe de processus très importante composée des différences de martingales.

Définition 6.
On dira qu'un processus stochastique $\{\mathbf{X}_t,\ t \in \mathbb{Z}\}$ est une différence de martingales s'il satisfait la condition : $\mathbf{E}\left(\mathbf{X}_t/\mathcal{F}_{t-1}\right) = 0$*, pour tout* $t \in \mathbb{N}$*, où* $\mathcal{F}_t = \sigma\left\{\mathbf{X}_t, \mathbf{X}_{t-1}, \ldots\right\}$ *désigne la tribu engendrée par* $\{\mathbf{X}_s,\ s \leq t\}$*.*

La classe des différences de martingales permet une plus grande variété que les bruits blancs forts dans les modèles de séries chronologiques. En effet, le bruit blanc fort est souvent considéré très restrictif. D'un autre côté, le bruit blanc faible est trop général. Les différences de martingales sont des bruits blancs faibles (bien que des bruits blancs faibles ne sont pas forcément des différences de martingales). Beaucoup de résultats théoriques valides pour les bruits blancs forts ont des équivalents pour des différences de martingales, sous certaines conditions. Nous ferons par exemple appel à un théorème limite central pour différences de martingales pour obtenir certains résultats de l'article. Voir par exemple Stout (1974).

1.2.1 Estimation de la moyenne et des autocovariances

Le vecteur moyenne et la matrice des variances-covariances définis ci-dessus sont rarement disponibles en pratique. Généralement, nous disposons d'une réalisation $\mathbf{X}_1, \ldots, \mathbf{X}_n$ du processus de taille n finie. Il est alors nécessaire de chercher à estimer la fonction d'autocovariance du processus sous-jacent. On peut estimer la moyenne du processus $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ par la moyenne échantillonale :

$$\bar{\mathbf{X}}_n = n^{-1} \sum_{t=1}^{n} \mathbf{X}_t.$$

De même, une version échantillonale de l'autocovariance $\boldsymbol{\Gamma}_X(h)$ est :

$$\hat{\boldsymbol{\Gamma}}_{\mathbf{X}}(h) = \begin{cases} n^{-1} \sum_{t=h+1}^{n} \left(\mathbf{X}_{t+h} - \bar{\mathbf{X}}_n\right)\left(\mathbf{X}_t - \bar{\mathbf{X}}_n\right)^\top, & si\ 0 \leq h \leq n-1, \\ \hat{\boldsymbol{\Gamma}}(-h)^\top, & si\ -n+1 \leq h < 0. \end{cases}$$

Les propriétés de ces estimateurs sont discutées dans Brockwell et Davis (1991) ou Fuller (1996), par exemple.

1.2.2 Modélisation des séries stationnaires

Construire le modèle sous-jacent est un enjeu majeur dans la démarche de modélisation d'un processus stochastique $\{\mathbf{X}_t, t \in \mathbb{Z}\}$. Nous introduisons dans cette partie une classe de modèles extrêmement importante et très utilisée pour la prévision des processus stationnaires : la classe des modèles autorégressifs-moyennes-mobiles multivariés, généralement connue sous l'acronyme VARMA. Elle est construite sur la base du bruit blanc défini dans la définition 5.

Définition 7.
Un processus stationnaire $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ centré en zéro, c'est à dire $E(\mathbf{X}_t) = \mathbf{0}$, de dimension d est un VARMA d'ordres p et q, noté VARMA(p,q), s'il vérifie la relation suivante :

$$\mathbf{X}_t - \mathbf{\Phi}_1 \mathbf{X}_{t-1} - \ldots - \mathbf{\Phi}_p \mathbf{X}_{t-p} = \mathbf{Z}_t + \mathbf{\Theta}_1 \mathbf{Z}_{t-1} + \ldots + \mathbf{\Theta}_q \mathbf{Z}_{t-q}, \tag{1.5}$$

où $\{\mathbf{Z}_t, t \in \mathbb{Z}\} \sim BB(\mathbf{0}, \mathbf{\Sigma})$, et $\mathbf{\Phi}_i,\ i = 1, \ldots, p$ et $\mathbf{\Theta}_j,\ j = 1, \ldots, q$ sont des matrices $d \times d$, avec $\mathbf{\Phi}_p \neq \mathbf{0}$ et $\mathbf{\Theta}_q \neq \mathbf{0}$.

L'écriture précédente peut être simplifiée de la manière suivante :

$$\mathbf{\Phi}(B)\mathbf{X}_t = \mathbf{\Theta}(B)\mathbf{Z}_t,\ \{\mathbf{Z}_t, t \in \mathbb{Z}\} \sim BB(\mathbf{0}, \mathbf{\Sigma}),$$

avec $\mathbf{\Phi}(z) = \mathbf{I} - \mathbf{\Phi}_1 z - \cdots - \mathbf{\Phi}_p z^p$ et $\mathbf{\Theta}(z) = \mathbf{I} + \mathbf{\Theta}_1 z + \cdots + \mathbf{\Theta}_q z^q$ qui sont les polynômes (matriciels) et $\mathbf{I}$ la matrice identité $d \times d$ et enfin B est l'opérateur retard défini par $BX_t = X_{t-1}$. Il arrive fréquemment qu'on restreigne la définion précédente en supposant que $\{\mathbf{Z}_t, t \in \mathbb{Z}\}$ est un bruit blanc fort ; on parle alors de processus VARMA fort. Cette restriction permet de simplifier l'étude lors de l'estimation. Les cas $\mathbf{\Theta}(z) = \mathbf{I}$ (resp $\mathbf{\Phi}(z) = \mathbf{I}$) représentent deux cas particuliers de processus très utiles en pratique, les processus vectoriels autorégressifs d'ordre p notés VAR(p) (les processus vectoriels moyennes-mobiles d'ordre q sont notés VMA(q)).

1.2.3 Notion de causalité et d'inversibilité d'un processus VARMA

Deux notions importantes sont généralement reliées aux processus VARMA. Il s'agit de la notion de causalité et d'inversibilité. Elles sont très utiles en ce sens qu'elles facilitent les calculs des prévisions linéaires optimales du processus. Avant de donner la définition des concepts, rappelons la définition suivante :

Définition 8.
Une suite de matrices $\{\mathbf{A}_i, i \in \mathbb{Z}\}$ *est absolument sommable si la limite suivante existe :*

$$lim_{n\to\infty} \sum_{i=-n}^{n} \|\mathbf{A}_i\| < \infty,$$

où $\|\cdot\|$ *dénote la norme euclidienne.*

Définition 9.
Un processus VARMA $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ *est dit causal, s'il existe une suite de matrices* $\{\mathbf{\Psi}_j\}$ *de coefficients absolument sommable telle que :*

$$\mathbf{X}_t = \sum_{j=0}^{\infty} \mathbf{\Psi}_j \mathbf{Z}_{t-j}, \text{ pour tout } t,$$

où $\{\mathbf{Z}_t, t \in \mathbb{Z}\} \sim BB(\mathbf{0}, \mathbf{\Sigma})$. *Pour un processus VARMA, cette notion est équivalente à la condition* $det\{\mathbf{\Phi}(z)\} \neq 0$ *pour tout* z *complexe tel que* $|z| \leq 1$ *où* $det\{\ldots\}$ *représente le déterminant.*

Les matrices $\{\mathbf{\Psi}_j\}$ sont calculées récursivement par l'équation :

$$\mathbf{\Psi}_j = \mathbf{\Theta}_j + \sum_{k=1}^{\infty} \mathbf{\Phi}_k \mathbf{\Psi}_{j-k}, \ j = 0, \ldots,$$

où $\mathbf{\Theta}_0 = \mathbf{I}$, $\mathbf{\Theta}_j = \mathbf{0}$ pour $j > q$, $\mathbf{\Phi}_j = \mathbf{0}$ pour $j > p$ et $\mathbf{\Psi}_j = \mathbf{0}$ pour $j < 0$. Voir par exemple Brockwell et Davis (1991, p. 242).

Définition 10.
Un processus VARMA $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ *est dit inversible, s'il existe une suite de matrices* $\{\mathbf{\Pi}_j\}$ *de coefficients absolument sommable telle que :*

$$\mathbf{Z}_t = \sum_{j=0}^{\infty} \mathbf{\Pi}_j \mathbf{X}_{t-j}, \text{ pour tout } t,$$

où $\{\mathbf{Z}_t, t \in \mathbb{Z}\} \sim BB(\mathbf{0}, \mathbf{\Sigma})$.

Cette nouvelle définition est équivalente à la condition $det\{\mathbf{\Theta}(z)\} \neq 0$ pour tout z complexe tel que $|z| \leq 1$. Il est aussi possible de calculer les matrices $\{\mathbf{\Pi}_j\}$ en utilisant de manière récursive l'équation :

$$\mathbf{\Pi}_j = -\mathbf{\Phi}_j - \sum_{k=1}^{\infty} \mathbf{\Theta}_k \mathbf{\Pi}_{j-k}, \ j = 0, \ldots,$$

avec $\mathbf{\Phi_0} = -\mathbf{I}$, $\mathbf{\Phi}_j = \mathbf{0}$ pour $j > p$, $\mathbf{\Theta}_j = \mathbf{0}$ pour $j > q$ et $\mathbf{\Pi}_j = \mathbf{0}$ pour $j < 0$. Voir Brockwell et Davis (1991, p. 242).

1.3 Processus conditionnellement hétéroscédastiques

Dans la section précédente, nous avons introduit la classe des modèles VARMA très répandue dans la modélisation de nombreux processus stationnaires multivariés lorsque le terme d'erreur est un bruit blanc fort. Or cette classe de processus s'avère très inadaptée pour la modélisation des séries financières. Par exemple, les séries des rendements d'un indice boursier présentent de nombreuses caractéristiques comme des effets de grappe de la volatilité (ou *volatility clustering*) ; les fortes variations tendent à être suivies par des variations importantes, alors que les variations de faible amplitude semblent également regroupées ensemble. Aussi, les distributions de probabilité empiriques des séries des rendements ne correspondent généralement pas à une distribution gaussienne. D'autres caractéristiques discutées dans Francq et Zakoian (2009) suggèrent que les modèles VARMA sont inappropriés pour la modélisation des séries financières. Dans le contexte des séries univariées, Engle (1982) a introduit la classe des modèles autorégressifs conditionnellement hétéroscédastique (ARCH), afin de prendre en compte le caractère aléatoire de la variance conditionnelle. Cette méthode a été généralisée par Bollerslev (1986), qui a introduit la classe des modèles ARCH généralisés (GARCH). Nous introduisons une classe de modèles qui permet de spécifier une structure particulière pour le bruit blanc d'un modèle VARMA.

Définition 11.
On dit que $\{\boldsymbol{\epsilon}_t, t \in \mathbb{Z}\}$ est un processus GARCH(p,q) multivarié, si :

$$\begin{cases} \boldsymbol{\epsilon}_t = \mathbf{L}_t \boldsymbol{\eta}_t, \\ \mathbf{V}_t = \mathbf{f}\left(\mathbf{V}_{t-1}, \ldots, \mathbf{V}_{t-p}, \boldsymbol{\epsilon}_{t-1}, \ldots, \boldsymbol{\epsilon}_{t-q}\right), \end{cases}$$

où $\mathbf{V}_t$ est une matrice supposée définie positive telle que $\mathbf{L}_t \mathbf{L}_t^\top = \mathbf{V}_t = Var\left(\boldsymbol{\epsilon}_t / \{\boldsymbol{\epsilon}_u, \ u < t\}\right)$ et $\{\boldsymbol{\eta}_t, t \in \mathbb{Z}\}$ est un processus de dimension d de moyenne nulle et constitué de vecteurs aléatoires indépendants et identiquement distribués. La fonction f est une transformation $\mathcal{F}_{t-1}$ mesurable.

La définition ci-dessus est générale et on note que le terme d'erreur dans la définition 11 satisfait les conditions d'une différence de martingales, c'est donc un terme d'erreur qui de manière générale ne sera pas fort (à moins que $\mathbf{V}_t \equiv \mathbf{V}$ soit une matrice constante). Plusieurs spécifications sont utilisées en pratique ; ces modèles diffèrent par les hypothèses faites sur les matrices $\mathbf{V}_t$. Nous présenterons deux des plus répandus, à savoir les modèles DVEC et BEKK. L'acronyme DVEC veut dire Diagonal VEC alors que BEKK provient des initiales des noms des auteurs qui ont proposé le modèle, c'est à dire Baba, Engle, Kraft et Kroner. Le modèle à été introduit par Engle et Kroner (1995). Ces modèles se trouvent également dans Zivot et Wang (2003).

1.3.1 Modèles DVEC

Ce modèle est un cas particulier de la formulation générale introduite plus haut, et a été étudié par Bollerslev, Engle et Wooldridge (1988). L'expression de $\mathbf{V}_t$ pour ce modèle est donnée par :

$$\mathbf{V}_t = \mathbf{A}_0 + \sum_{i=1}^{p} \mathbf{A}_i \odot \left(\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}_{t-i}^\top\right) + \sum_{j=1}^{q} \mathbf{B}_j \odot \mathbf{V}_{t-j},$$

où $'\odot'$ est le produit d'Hadamard, c'est-à-dire que le produit des matrices $\mathbf{A}$ et $\mathbf{B}$ se fait élément par élément : $\mathbf{A} \odot \mathbf{B} = (a_{ij}b_{ij})$, où $\mathbf{A}$ et $\mathbf{B}$ sont compatibles. Une matrice des variances-covariances se doit d'être symétrique et définie positive. Pour assurer la symétrie des matrices V_t spécifiées ci-dessus, il suffit de supposer que les matrices $\mathbf{A}_0,\ \mathbf{A}_i,\ \mathbf{B}_j,\ i = 1,\ldots,p,\ j = 1,\ldots,q$ sont symétriques. Cependant rien ne peut garantir que ces matrices sont définies positives. Il existe alors d'autres formulations des matrices des variances-covariances permettant de garantir que celles-ci soient symétriques et définies positives. Notons qu'une condition suffisante garantissant que les matrices $\mathbf{V}_t$ sont définies positives est que les matrices $\mathbf{A}_0,\ \mathbf{A}_i,\ \mathbf{B}_j,\ i = 1,\ldots,p,\ j = 1,\ldots,q$ soient définies positives. Ding (1994) et Bollerslev, Engle et Nelson (1994) ont alors proposé le modèle suivant :

$$\mathbf{V}_t = \mathbf{A}_0\mathbf{A}_0^\top + \sum_{i=1}^{p} \left(\mathbf{A}_i\mathbf{A}_i^\top\right) \odot \left(\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}_{t-i}^\top\right) + \sum_{j=1}^{q} \left(\mathbf{B}_j\mathbf{B}_j^\top\right) \odot \mathbf{V}_{t-j}.$$

Une autre variante de ce modèle consiste à remplacer les matrices $\mathbf{A}_0,\ \mathbf{A}_i,\ \mathbf{B}_j,\ i = 1,\ldots,p,\ j = 1,\ldots,q$ par des scalaires $a_0,\ a_i,\ b_j,\ i = 1,\ldots,p,\ j = 1,\ldots,q$, permettant ainsi de faciliter l'inférence.

1.3.2 Modèles BEKK

Le modèle BEKK est une classe de modèles encore plus générale et moins restrictive que les modèles DVEC. Initié par Engle et Kroner (1995), la variance conditionnelle est formulée de la façon suivante :

$$\mathbf{V}_t = \mathbf{A}_0\mathbf{A}_0^\top + \sum_{i=1}^{p} \mathbf{A}_i \odot \left(\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}_{t-i}^\top\right) \mathbf{A}_i^\top + \sum_{j=1}^{q} \mathbf{B}_j \odot \mathbf{V}_{t-j}\mathbf{B}_j^\top,$$

où $\mathbf{A}_0$ est une matrice triangulaire inférieure, mais les matrices $\mathbf{A}_i,\ i = 1,\ldots,p,$ et $\mathbf{B}_j,\ j = 1,\ldots,q$, sont des matrices quelconques.

Le lecteur intéressé peut consulter les détails des modèles DVEC et BEKK dans le livre de Zivot et Wang (2003).

1.4 Notion de causalité entre deux séries chronologiques

La notion de causalité est importante dans la littérature statistique et économétrique. Dans beaucoup de situations, il est souvent important d'étudier les relations de cause à effet entre des séries chronologiques. En économie et finance par exemple, si les informations arrivent en grappes, les rendements des actifs ou les prix peuvent présenter une volatilité fonction du temps. L'étude de la causalité peut dans ce cas nous aider à comprendre comment le marché ou les actifs s'adaptent à l'arrivée de nouvelles informations.

Dans cette section, nous présenterons les deux concepts principaux entourant la notion de causalité, à savoir la causalité en moyenne et en variance. Soit $\{\mathbf{X}_t,\ t \in \mathbb{Z}\}$ un processus de dimension d ayant les composantes $\mathbf{X}_{1t}$ et $\mathbf{X}_{2t}$ de dimensions d_1 et d_2 respectivement, avec $d = d_1 + d_2$. Posons $\mathcal{F}_{it},\ i = 1, 2$, l'information disponible sur les séries $\mathbf{X}_{it}$ jusqu'à l'instant t, et $\mathcal{F}_t = (\mathcal{F}_{1t}, \mathcal{F}_{2t})$. Définissons $\overline{\mathbf{X}_{1t}}$ l'espace de Hilbert engendré par les variables $\{\mathbf{X}_{1s},\ s \leq t\}$. On note également par $\mathbf{P}\left\{\mathbf{X}_{i,t}/\overline{\mathbf{X}_t}\right\}$ le vecteur $\left(\mathbf{P}\left\{\mathbf{X}_{i1,t}/\overline{\mathbf{X}_t}\right\},\ldots,\mathbf{P}\left\{\mathbf{X}_{id_i,t}/\overline{\mathbf{X}_t}\right\}\right)^\top$. Rappellons que $\mathbf{P}\left\{\mathbf{X}_{ij,t}/\overline{\mathbf{X}_t}\right\}$ est la projection orthogonale de $\mathbf{X}_{ij,t}$ sur $\overline{\mathbf{X}_t}$.

1.4.1 Causalité en moyenne

Nous présentons ci-dessous deux définitions de la causalité en moyenne au sens de Granger. Les définitions et résultats de cette section proviennent de Comte et Lieberman (2000). En effet, la causalité en variance dans un cadre multivarié peut être attribuée à Comte et

Lieberman (2000). Ces auteurs ont proposé deux définitions pour la causalité en variance. Le premier type est très similaire au concept de causalité au sens de Granger, alors que le second type est un concept de causalité au sens de Granger linéaire utilisant les projections et les espaces de Hilbert. On note que le second type est correctement adapté afin de cerner la causalité dans le second moment.

Définition 12.
On dit que $\mathbf{X}_2$ *ne cause pas* $\mathbf{X}_1$ *au sens de Granger et on dénote* $\mathbf{X}_2 \overset{G}{\nrightarrow} \mathbf{X}_1$*, si*

$$E\left\{\mathbf{X}_{1,t+1}/\mathcal{F}_t\right\} = E\left\{\mathbf{X}_{1,t+1}/\mathcal{F}_{1t}\right\}, \ \forall t \in \mathbb{Z}. \tag{1.6}$$

On dit que $\mathbf{X}_2$ *ne cause pas* $\mathbf{X}_1$ *linéairement au sens de Granger et on note* $\mathbf{X}_2 \overset{LG}{\nrightarrow} \mathbf{X}_1$*, si*

$$P\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}}_t\right\} = P\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}}_{1t}\right\}, \ \forall t \in \mathbb{Z}. \tag{1.7}$$

La définition 12 est ainsi le concept de causalité en moyenne, que nous présentons ici à titre de rappel. Nous aurons également besoin de tester la causalité en moyenne avant la causalité en variance, comme dans Cheung et Ng (1996). Ainsi, selon (1.6), $\mathbf{X}_2$ ne cause pas $\mathbf{X}_1$ au sens de Granger implique que la meilleure prévision est identique, si on conditionne sur le passé de $\mathbf{X}_1$ et $\mathbf{X}_2$, ou si on conditionne que sur le passé de $\mathbf{X}_1$. Ainsi, si $\mathbf{X}_2$ ne cause pas $\mathbf{X}_1$, la connaissance du passé de $\mathbf{X}_2$ n'améliore pas la meilleure prévision qu'est l'espérance conditionnelle. La définition (1.7) s'interprète similairement, où l'on se restreint cependant au choix de la meilleure prévision linéaire. Dans le cas où le processus $\{\mathbf{X}_t, \ t \in \mathbb{Z}\}$ est un processus gaussien, les énoncés (1.6) et (1.7) dans la définition 12 sont équivalentes. En effet $E\left\{\mathbf{X}_{t+1}/\mathcal{F}_t\right\}$ représente le meilleur prédicteur alors que $P\left\{\mathbf{X}_{t+1}/\overline{\mathbf{X}}_t\right\}$ représente le meilleur prédicteur linéaire. Ces deux prédicteurs qui en général sont différents, coïncident dans le cas d'un processus gaussien.

La notion de causalité en moyenne ne suffit généralement pas pour décrire les processus stochastiques dont les matrices des variances-covariances conditionnelles varient avec le temps. C'est ainsi que les définitions précédentes ont été étendues de manière naturelle à la notion de causalité en variance. Avant de présenter la définition formelle, il est nécessaire de préciser certaines notations supplémentaires pour des raisons de cohésion. Soit $\overline{\mathbf{X}_t^2}$ l'espace de Hilbert engendré par les variables $X_{ij,\,s} \cdot X_{kl,\,s}, \ s \leq t, \ 1 \leq i, \ k \leq 2, \ 1 \leq j \leq d_i, \ 1 \leq k \leq d_k$ et soit $\overline{\mathbf{X}_{1t}^2}$ l'espace de Hilbert engengré par $X_{1j,s} \cdot X_{1l,s}, \ 1 \leq j, \ l \leq d_1$. Ces extensions sont nécessaires quand l'on veut décrire la causalité en variance. Voir aussi Comte et Lieberman (2000, p. 537).

1.4.2 Causalité en variance

Le rappel sur la causalité en moyenne qui a fait l'objet de la définition 12 devrait s'avérer utile dans la motivation d'une définition pour la causalité en variance. Ceci fait l'objet de la définition 13.

Définition 13.
On dit que $\mathbf{X}_2$ *ne cause pas* $\mathbf{X}_1$ *en variance et on note* $\mathbf{X}_2 \overset{G_V}{\nrightarrow} \mathbf{X}_1$*, si*

$$V\left\{\mathbf{X}_{1,t+1}/\mathcal{F}_t\right\} = V\left\{\mathbf{X}_{1,t+1}/\mathcal{F}_{1t}\right\}, \ \forall t \in \mathbb{Z}. \tag{1.8}$$

On dit que $\mathbf{X}_2$ *ne cause pas* $\mathbf{X}_1$ *linéairement en variance et on note* $\mathbf{X}_2 \overset{LG_V}{\nrightarrow} \mathbf{X}_1$*, si* $\forall t \in \mathbb{Z}$*,*

$$\begin{aligned} &P\left[\left(\mathbf{X}_{1,t+1} - \mathbf{P}\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}_t}\right\}\right)\left(\mathbf{X}_{1,t+1} - \mathbf{P}\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}_t}\right\}\right)^\top/\overline{\mathbf{X}_t^2}\right] = \\ &P\left[\left(\mathbf{X}_{1,t+1} - \mathbf{P}\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}_{1t}}\right\}\right)\left(\mathbf{X}_{1,t+1} - \mathbf{P}\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}_{1t}}\right\}\right)^\top/\overline{\mathbf{X}_{1t}^2}\right]. \end{aligned} \tag{1.9}$$

La définition 13 prend tout son sens lorsque l'on interprète la variance conditionnelle comme une mesure de volatilité. Ainsi la volatilité reste inchangée lorsque l'on conditionne sur le passé de $\mathbf{X}_1$ et de $\mathbf{X}_2$, versus conditionner uniquement sur $\mathbf{X}_1$: on dira alors que $\mathbf{X}_2$ ne cause pas $\mathbf{X}_1$ en variance. Ceci correspond à la définition 13, équation (1.8). Une autre notion de causalité très liée aux concepts précédents de causalité en moyenne et en variance, est la causalité au second ordre. Elle a été introduite par Granger *et al.* (1986), et est très utile dans la décomposition de la causalité en variance. En effet, il est possible de caractériser l'absence de causalité en variance si et seulement s'il y a à la fois absence de causalité en moyenne et absence de causalité au second ordre.

Définition 14.
On dit que $\mathbf{X}_2$ *ne cause pas* $\mathbf{X}_1$ *au second ordre et on note* $\mathbf{X}_2 \overset{G_2}{\nrightarrow} \mathbf{X}_1$*, si* $\forall t \in \mathbb{Z}$*,*

$$\begin{aligned} &E\left[\left(\mathbf{X}_{1,t+1} - E\left\{\mathbf{X}_{1,t+1}/\mathcal{F}_t\right\}\right)\left(\mathbf{X}_{1,t+1} - E\left\{\mathbf{X}_{1,t+1}/\mathcal{F}_t\right\}\right)^\top/\mathcal{F}_t\right] = \\ &E\left[\left(\mathbf{X}_{1,t+1} - E\left\{\mathbf{X}_{1,t+1}/\mathcal{F}_t\right\}\right)\left(\mathbf{X}_{1,t+1} - E\left\{\mathbf{X}_{1,t+1}/\mathcal{F}_t\right\}\right)^\top/\mathcal{F}_{1t}\right]. \end{aligned} \tag{1.10}$$

On dit que $\mathbf{X}_2$ *ne cause pas* $\mathbf{X}_1$ *linéairement au second ordre et on note* $\mathbf{X}_2 \overset{LG_2}{\nrightarrow} \mathbf{X}_1$*, si* $\forall t \in \mathbb{Z}$*,*

$$\begin{aligned} &P\left[\left(\mathbf{X}_{1,t+1} - \mathbf{P}\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}_t}\right\}\right)\left(\mathbf{X}_{1,t+1} - \mathbf{P}\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}_t}\right\}\right)^\top/\overline{\mathbf{X}_t^2}\right] = \\ &P\left[\left(\mathbf{X}_{1,t+1} - \mathbf{P}\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}_t}\right\}\right)\left(\mathbf{X}_{1,t+1} - \mathbf{P}\left\{\mathbf{X}_{1,t+1}/\overline{\mathbf{X}_t}\right\}\right)^\top/\overline{\mathbf{X}_{1t}^2}\right]. \end{aligned} \tag{1.11}$$

Nous allons clore cette section en énonçant et sans démonstration une proposition présentant les relations d'équivalences qui existent entre les différents concepts de causalité définis plus haut. Pour plus de précision sur ces notions, on peut consulter Comte et Lieberman (2000).

Proposition 1.

$$\mathbf{X}_2 \overset{G_V}{\nrightarrow} \mathbf{X}_1 \Leftrightarrow \left(\mathbf{X}_2 \overset{G}{\nrightarrow} \mathbf{X}_1 \ et\ \mathbf{X}_2 \overset{G_2}{\nrightarrow} \mathbf{X}_1\right).$$

$$\mathbf{X}_2 \overset{LG_V}{\nrightarrow} \mathbf{X}_1 \Leftrightarrow \left(\mathbf{X}_2 \overset{LG}{\nrightarrow} \mathbf{X}_1 \ et\ \mathbf{X}_2 \overset{LG_2}{\nrightarrow} \mathbf{X}_1\right).$$

Il découle de cette proposition que l'existence de causalité en moyenne entraîne systématiquement l'existance de causalité en variance. Cependant on ne peut rien dire sur la causalité au second ordre du simple fait d'existence de la causalité en moyenne. Une contribution significative de Comte et Lieberman (2000) est que l'absence de causalité au second ordre peut être testée comme des restrictions sur des paramètres d'une classe générale de modèles VARMA avec erreurs GARCH. Dans l'article qui suit, la proposition 1 sera très utilile car elle suggérera une stratégie afin de tester la causalité en variance. Dans un premier temps, nous allons tester avec des outils disponibles dans la littérature la causalité en moyenne. Par la suite, nous allons développer des statistiques de test pour tester la causalité au second ordre.

Chapitre 2

Article

Ce chapitre contient l'article dans son intégralité intitulé « On testing for causality in variance between two multivariate time series » soumis dans une revue scientifique avec comité de lecture le 27 octobre 2010. Le premier auteur est également l'auteur du présent mémoire.

On testing for causality in variance between two multivariate time series*

Herbert NKWIMI†

Département de mathématiques et statistique

Université de Montréal

Pierre DUCHESNE

Département de mathématiques et statistique

Université de Montréal

October 27, 2010

Abstract

Multivariate conditional heteroscedasticity models have been found useful for modelling financial data. In many applications, to verify the existence of a relationship between two multivariate time series represents an important consideration. In this article, the procedure developed by Cheung and Ng (1996) designed to test causality in variance for univariate time series is generalized in several directions. Relying on the work of El Himdi and Roy (1997) and Duchesne (2004), a first approach proposes test statistics based on residual cross-covariance matrices of squared (standardized) residuals and cross products of (standardized) residuals. The new test statistics converge in distribution toward convenient chi-square distributions under the null hypothesis of absence of causality in variance. In a second approach, transformed residuals are defined as in Ling and Li (1997) for each residual vector time series, and test statistics are constructed based on the cross-correlations of these transformed residuals. In both approaches, test statistics at individual lags are developed, and also portmanteau type test statistics; both procedures reduce to Cheung and Ng's test statistics in the case of two univariate time series. The proposed methodology can be used to determine the directions of causality in variance, and appropriate test statistics are presented. Monte Carlo

*Abbreviated title: "Causality in variance between two multivariate time series".

Corresponding author: Pierre Duchesne, Université de Montréal; Département de mathématiques et de statistique; C.P. 6128 Succursale Centre-Ville; Montréal, Québec H3C 3J7; Canada.

†This work was supported by a grant from the Natural Science and Engineering Research Council of Canada.

simulation results show that the new test statistics offer satisfactory empirical properties. An application with real data illustrates the methods.
Key words and phrases: Causality in variance; Conditional heteroscedasticity; Portmanteau test statistics; Residual cross-correlations; Multivariate time series.

1. INTRODUCTION

The establishment of relationships between time series represents an important problem in analyzing real data. For example, in econometrics and finance, a natural question concerns the causation in conditional variance across financial asset price movements, whose study may contribute to understand several characteristics of the system under study. The financial literature suggests that changes of variance are closely related to the arrival of information, and thus practitioners are interested to describe how the market evaluates and assimilates the new information between the markets (see, e.g., Ross (1989) and Cheung and Ng (1990, 1996)). From a financial perspective, the absence of causality in variance suggests that the principal sources of disturbances are changes in asset- or market-specific fundamentals (see, e.g., Hong (2001)). On the other hand, in the presence of causality in variance, an important change in volatility in a given market may potentially increases the volatility in other markets as well. Appropriate procedures to detect causality in variance are necessary, but it is also important to have simple methods to identify the causation patterns, which provides useful insight concerning the characteristics and dynamics of economic and financial prices (see, e.g., Cheung and Ng (1996)). From a statistical point of view, the study of that kind of relations is also of importance in building time series models at the model identification stage. In this paper, we are concerned with portfolios and relationships between two multivariate models. Multivariate time series models with general specifications for the conditional variance may depend on a large number of independent parameters. Thus, it appears critical to have methods which are simple to use and to implement with computer software, whose results can be interpreted easily.

In order to test for causality in variance between two univariate time series, $\{X_{1t}, t \in \mathbb{Z}\}$ and $\{X_{2t}, t \in \mathbb{Z}\}$, Cheung and Ng (1996) developed a two-stage procedure, which is asymptotically robust to distributional assumptions. In the first stage of their procedure, univariate time series models are fitted to each time series. Possible specifications include autoregressive-moving-average (ARMA) models and (generalized) autoregressive conditional heteroscedastic [(G)ARCH] processes. In the second stage, the resulting series of squared standardized residuals, denoted $\{\hat{u}_{1t}, t = 1, \ldots, n\}$ and $\{\hat{u}_{2t}, t = 1, \ldots, n\}$, are calculated, where n represents the sample size. The test statistics are then based on the residual cross-correlations at lags k, $|k| < n$, defined as:

$$r_{\hat{u}_1\hat{u}_2}(k) = c_{\hat{u}_1\hat{u}_2}(k)\{c_{\hat{u}_1\hat{u}_1}(0)c_{\hat{u}_2\hat{u}_2}(0)\}^{-1/2},$$

where $c_{\hat{u}_1\hat{u}_2}(k)$ represents the kth lag residual cross-covariance given by:

$$c_{\hat{u}_1\hat{u}_2}(k) = n^{-1}\sum_{t=k+1}^{n}\{\hat{u}_{1t} - \text{ave}(\hat{u}_1)\}\{\hat{u}_{2,t-k} - \text{ave}(\hat{u}_2)\},\ k = 0,\ldots,n-1,$$

and $c_{\hat{u}_1\hat{u}_2}(k) = c_{\hat{u}_2\hat{u}_1}(-k)$ when $k = -1,\ldots,-n+1$, with $\text{ave}(\hat{u}_1) = n^{-1}\sum_{t=1}^{n}\hat{u}_{1t}$ and $\text{ave}(\hat{u}_2) = n^{-1}\sum_{t=1}^{n}\hat{u}_{2t}$. Consider a finite set of lags $k_1, k_2, \ldots, k_m$. Subject to certain conditions on the moments of the innovation processes, Cheung and Ng (1996) established, using results of Haugh (1976), that the random vector $n^{1/2}\left(r_{\hat{u}_1\hat{u}_2}(k_1),\ldots,r_{\hat{u}_1\hat{u}_2}(k_m)\right)^\top$ converges in distribution to a m-dimensional multinormal random vector $\mathcal{N}_m(\mathbf{0},\mathbf{I}_m)$, with $\mathbf{I}_m$ the $m\times m$ identity matrix, under the null hypothesis of no causality in variance between the two time series $\{X_{1t}\}$ and $\{X_{2t}\}$. As applications of that result, several useful test statistics having asymptotic chi-square distributions can be easily constructed. First, one-lag test statistics may rely on $\text{CN}_n(k) = nr^2_{\hat{u}_1\hat{u}_2}(k)$, which has asymptotically a chi-square distribution with one degree of freedom, that is $\text{CN}_n(k) \xrightarrow{d} \chi^2_1$, where χ^2_d represents the chi-square distribution with d degrees of freedom. Alternatively, a portmanteau test statistic can be defined:

$$\text{CN}_n(L:U) = n\sum_{k=L}^{U} r^2_{\hat{u}_1\hat{u}_2}(k), \tag{1}$$

with $L \leq U$, which converges in distribution toward a chi-square distribution with $U-L+1$ degrees of freedom, that is $\text{CN}_n(L:U) \xrightarrow{d} \chi^2_{U-L+1}$. The choices of the lower and upper bounds L and U depend on the specification of the alternative hypothesis. For example, the choice $(L,U) = (-M,M)$, where M is a maximal lag order satisfying $M < n$, can be used when there is no information on the direction of the causality in variance. The choice $(L,U) = (-M,-1)$ represents a natural choice to test the null hypothesis of no causality in variance against causality in variance from $\{X_{1t}\}$ to $\{X_{2t}\}$. Similarly, $(L,U) = (1,M)$ is appropriate for testing causality in variance from $\{X_{2t}\}$ to $\{X_{1t}\}$. See Cheung and Ng (1986) for additional details on these descriptive causality methods. El Himdi and Roy (1997) advocated similar procedures when testing for Granger causality (in mean) between two multivariate time series. For testing causality in mean, they found these kinds of methods both simple to use and to execute, since there is no need to estimate a global model for the two series. They applied successfully these methods to detect the causation patterns in the conditional means between real economic data.

The problem of causality in variance between two univariate time series has been also considered in Wong and Li (1996), who derive the asymptotic distribution of the cross-correlations of squared residuals in integrated ARMA models when the innovations are M-dependent. Their results have been generalized in Wong and Li (2002). A Lagrange multiplier test statistic for causality in variance between two univariate time series using multivariate GARCH models is studied in Hafner and Herwartz (2006) and these authors also develop Wald's test statistics for causality in variance between two univariate time series using multivariate models, see Hafner and Herwartz (2008).

All the preceding literature proposed methods for detecting causality in variance between univariate time series. In a multivariate setting, Comte and Lieberman (2000) provided two definitions of variance noncausality; a Granger-type noncausality and a linear Granger noncausality through projections on Hilbert spaces. While Comte and Lieberman (2000) considered VARMA-GARCH models, Caporin (2007) extended their results for multivariate GARCH-in-mean models. Here, the main objective of this article is to generalize for two multivariate time series, noted $\mathbf{X}_1 = \{\mathbf{X}_{1t}, t \in \mathbb{Z}\}$ and $\mathbf{X}_2 = \{\mathbf{X}_{2t}, t \in \mathbb{Z}\}$, the approach of Cheung and Ng (1996) for testing causality in variance between the two series. Two possible generalizations are considered. In the first approach, using results of El Himdi and Roy (1997), and using the matricial measures of dependence introduced in Duchesne (2004), we obtain the asymptotic distributions of these residual cross-covariance matrices under the null hypothesis of no causality in variance. New test statistics are proposed as applications of that result. In the second approach, we define transformed residuals for each time series, using the method of Ling and Li (1997), and we derive the asymptotic distributions of the cross-correlations between the two transformed series. New test statistics are proposed, relying on squared transformed residuals cross-correlations. In both approaches, it is possible to define test statistics at a particular lag, and portmanteau type test statistics. These two classes of test statistics generalize Cheung and Ng's method, and in fact the generalized test statistics under both approaches reduce to Cheung and Ng's portmanteau test statistics when both time series are univariate.

The rest of the paper is organized as follows. In Section 2, some preliminaries are introduced, and we obtain the asymptotic distribution of a vector of lag-k residual cross-covariance matrices, using the dependence measure introduced in Duchesne (2004), under the null hypothesis of no causality in variance. We also discuss the asymptotic distribution of lag-k transformed residual cross-correlations, using the approach of Ling and Li (1997). As applications of these asymptotic results, we describe in Section 3 one-lag test statistics and portmanteau test statistics. We also discuss how to adapt the test statistics to detect the directions of causality in variance. In Section 4, we report the results of a small simulation experiment in order to study the finite-sample properties of the test statistics with respect to exact levels and powers. In Section 5, causality relations are investigated between two bivariate financial time series. The data set consists of two indexes coming from the North

American market, and two indexes of the European market. Section 6 offers concluding remarks.

2. PRELIMINARIES

The concept of causality in variance has been introduced in Granger, Robins and Engle (1986). Following their seminal work, Comte and Lieberman (2000) provided two new definitions, and in their framework non causality in variance exists if and only if there exists both non causality in mean and second-order non causality.

Consider two stationary and ergodic multivariate time series $\mathbf{X}_i = \{\mathbf{X}_{it}, t \in \mathbb{Z}\}$, such that $\mathbf{X}_{it} = (X_{it}(1), \ldots, X_{it}(d_i))^\top$, $i = 1, 2$. Let $\mathcal{F}_{it}$ denotes the information set generated by all past observations up to and including time t, that is $\mathcal{F}_{it} = \sigma\{\mathbf{X}_{is}, s \leq t\}$. Let $\mathcal{F}_t = (\mathcal{F}_{1t}, \mathcal{F}_{2t})$. According to the second-order non causality criterion, $\mathbf{X}_2$ does not cause $\mathbf{X}_1$ if the following relation is satisfied:

$$E\left[\{\mathbf{X}_{1,t+1} - E(\mathbf{X}_{1,t+1}|\mathcal{F}_t)\}\{\mathbf{X}_{1,t+1} - E(\mathbf{X}_{1,t+1}|\mathcal{F}_t)\}^\top|\mathcal{F}_{1t}\right] = V(\mathbf{X}_{1,t+1}|\mathcal{F}_t),\ \forall t \in \mathbb{Z},$$

noted $\mathbf{X}_2 \overset{G_2}{\nrightarrow} \mathbf{X}_1$. Using the definitions of Comte and Lieberman (2000), $\mathbf{X}_2$ does not cause $\mathbf{X}_1$ in variance if $V(\mathbf{X}_{1,t+1}|\mathcal{F}_{1t}) = V(\mathbf{X}_{1,t+1}|\mathcal{F}_t)$, $\forall t \in \mathbb{Z}$, denoted $\mathbf{X}_2 \overset{G_V}{\nrightarrow} \mathbf{X}_1$. In fact, it may be shown that $\mathbf{X}_2$ does not cause $\mathbf{X}_1$ in mean and $\mathbf{X}_2 \overset{G_2}{\nrightarrow} \mathbf{X}_1$ if and only if $\mathbf{X}_2 \overset{G_V}{\nrightarrow} \mathbf{X}_1$ (see Comte and Lieberman (2000, Proposition 1, p. 538)). In this paper, we are mainly concerned with the null hypothesis:

$$H_0 : \mathbf{X}_2 \overset{G_2}{\nrightarrow} \mathbf{X}_1, \tag{2}$$

and under the alternative hypothesis $E\left[\{\mathbf{X}_{1,t+1} - E(\mathbf{X}_{1,t+1}|\mathcal{F}_t)\}\{\mathbf{X}_{1,t+1} - E(\mathbf{X}_{1,t+1}|\mathcal{F}_t)\}^\top|\mathcal{F}_{1t}\right] \neq V(\mathbf{X}_{1,t+1}|\mathcal{F}_t)$. Concerning possible causality in mean, we note that it has been filtered out in defining the null hypothesis H_0. In practice, a sequential testing approach scheme may be implemented, checking first causality in mean. Then, if there is no relation between certain models, we may want to test for second-order non causality in a second step (see, e.g., Caporin (2007)). Cheung and Ng (1996) discussed the effects of causality in mean on tests for causality in variance and the interaction between these procedures. In their empirical example, they provided a model building strategy where augmented models were fitted before testing for causality in variance. In Section 5, we present a similar procedure where causality in mean is first studied using the methods and results of El Himdi and Roy (1997), applied however on properly standardized residuals.

The framework described in the following will allow us to develop test statistics for checking the null hypothesis (2). Suppose that $\{\mathbf{X}_{it}\}$, $i = 1, 2$, satisfy the following equations:

$$\mathbf{X}_{it} = \boldsymbol{\mu}_{it} + \boldsymbol{\epsilon}_{it}, \tag{3}$$

$$\boldsymbol{\epsilon}_{it} = \mathbf{L}_{it}\boldsymbol{\eta}_{it}, \; i = 1, 2. \tag{4}$$

In (4), the stochastic processes $\boldsymbol{\eta}_i = \{\boldsymbol{\eta}_{it}, t \in \mathbb{Z}\}$ are composed of independent and identically distributed random vectors with mean $\mathbf{0}$ and variance $\mathbf{I}_{d_i}$, $i = 1, 2$, respectively, and $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ are supposed to be independent under the null hypothesis. The conditional means and variances of $\{\mathbf{X}_{it}\}$, $i = 1, 2$, are given by $\boldsymbol{\mu}_{it}$ and $\mathbf{V}_{it} = \mathbf{L}_{it}\mathbf{L}_{it}^{\top}$, respectively. The conditional variance matrices $\mathbf{V}_{it}$, $i = 1, 2$, are supposed to be positive definite. We assume that the functions $\boldsymbol{\mu}_{it}$ and $\mathbf{L}_{it}$, measurable with respect to $\mathcal{F}_{i,t-1}$, rely on the model parameters $\boldsymbol{\beta}_i$, $\boldsymbol{\omega}_i$, and $\boldsymbol{\Lambda}_i$, such that $\boldsymbol{\mu}_{it} = \boldsymbol{\mu}_{it}(\boldsymbol{\beta}_i)$, $\mathbf{L}_{it} = \mathbf{L}_{it}(\boldsymbol{\omega}_i, \boldsymbol{\Lambda}_i)$, $i = 1, 2$. The dimensions of the model parameters $\boldsymbol{\beta}_i$, $\boldsymbol{\omega}_i$ and $\boldsymbol{\Lambda}_i$, $i = 1, 2$, are $p_i \times 1$, $d_i \times 1$ and $d_i \times P_i$, respectively. They are collected in two vectors $\boldsymbol{\theta}_i = (\boldsymbol{\beta}_i^{\top}, \boldsymbol{\omega}_i^{\top}, \mathrm{vec}^{\top}(\boldsymbol{\Lambda}_i))^{\top}$, $i = 1, 2$, where $\mathrm{vec}(\mathbf{A})$ represents the vector obtained by stacking the columns of $\mathbf{A}$. We suppose that $\boldsymbol{\mu}_{it}$, $\mathbf{V}_{it}$ and $\mathbf{L}_{it}$, $i = 1, 2$, are twice continuously differentiable with respect to $\boldsymbol{\theta}_i$ in a certain open neighborhood of $\boldsymbol{\theta}_{i0}$, almost surely.

Each model is fitted separately. We suppose that the estimators satisfy the following condition:

$$\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_{i0} = \mathbf{O}_P(n^{-1/2}), \; i = 1, 2. \tag{5}$$

Possible estimators are the Quasi Maximum Likelihood Estimators (QMLE), which have been studied in various multivariate models. Comte and Lieberman (2003) established the consistency and asymptotic normality of the QMLE estimators for the BEKK class of multivariate models. Asymptotic results are also given in Ling and McAleer (2003) for the Constant Conditional Correlation (CCC) formulation of an ARMA-GARCH model. See also Hafner and Preminger (2009) for the so-called Vec model. Recently, Francq and Zakoian (2010) provide asymptotic results for the CCC-GARCH(p,q) model under conditions which parallel those used in the univariate setting.

Having estimated each model, the residual cross-correlation measures are calculated. For each time series, we define properly standardized residuals:

$$\begin{aligned} \hat{\boldsymbol{\eta}}_{it} &= \hat{\mathbf{L}}_{it}^{-1}\hat{\boldsymbol{\epsilon}}_{it}, \\ &= \hat{\mathbf{L}}_{it}^{-1}(\mathbf{X}_{it} - \hat{\boldsymbol{\mu}}_{it}), \; i = 1, 2, \end{aligned}$$

which reduce to the usual standardized residuals when $d_1 = d_2 = 1$, that is $\hat{\eta}_{it} = \hat{\epsilon}_{it}/\hat{h}_{it}$, where $\hat{h}_{it}^2$ corresponds to the estimated conditional variance (see, e.g., Li and Mak (1994)).

We also calculate the residual cross-covariance matrices:

$$\mathbf{C}_{\hat{\boldsymbol{\eta}}_1\hat{\boldsymbol{\eta}}_2}(k) = \begin{cases} n^{-1}\sum_{t=k+1}^{n}\{\hat{\boldsymbol{\eta}}_{1t} - \mathrm{ave}(\hat{\boldsymbol{\eta}}_1)\}\{\hat{\boldsymbol{\eta}}_{2,t-k} - \mathrm{ave}(\hat{\boldsymbol{\eta}}_2)\}^\top, & k = 0, 1, ..., n-1, \\ \mathbf{C}^\top_{\hat{\boldsymbol{\eta}}_2\hat{\boldsymbol{\eta}}_1}(-k), & k = -n+1, ..., -1, \end{cases}$$

where $\mathrm{ave}(\hat{\boldsymbol{\eta}}_i) = n^{-1}\sum_{t=1}^{n}\hat{\boldsymbol{\eta}}_{it}$, $i = 1, 2$. These measures represent useful descriptive tools to investigate causality in mean, using for example the test statistics introduced in El Himdi and Roy (1997).

In order to detect second-order causality, an approach consists to study squared (standardized) residuals and cross products of (standardized) residuals. This leads us to introduce the random vectors:

$$\mathbf{Z}_{it} = \mathrm{vech}\{\mathbf{L}_{it}^{-1}\boldsymbol{\epsilon}_{it}\boldsymbol{\epsilon}_{it}^\top(\mathbf{L}_{it}^{-1})^\top\},\ i = 1, 2,$$

which are of dimension $d_i' = d_i(d_i + 1)/2$, $i = 1, 2$. It is assumed that $E(||\mathbf{Z}_{it}||^4) < \infty$, where $||\cdot||$ corresponds to the Euclidian norm. Note that for univariate processes, this reduces to the assumption made by Cheung and Ng (1996), see also Hong (2001, Assumption A.1). The half-vec operator $\mathrm{vech}(\cdot)$ is studied in Harville (1997, Chapter 16.4); for a $d \times d$ matrix $\mathbf{A}$, the vector $\mathrm{vech}(\mathbf{A})$ is composed of the $d(d+1)/2$ elements which are on and below the diagonal of $\mathbf{A}$. Based on the residuals of each time series, we define similarly $\hat{\mathbf{Z}}_{it} = \mathrm{vech}\{\hat{\mathbf{L}}_{it}^{-1}\hat{\boldsymbol{\epsilon}}_{it}\hat{\boldsymbol{\epsilon}}_{it}^\top(\hat{\mathbf{L}}_{it}^{-1})^\top\}$, $i = 1, 2$. We define $\boldsymbol{\Gamma}_{\mathbf{Z}_i\mathbf{Z}_i}(k) = \mathrm{cov}(\mathbf{Z}_{it}, \mathbf{Z}_{i,t-k})$ and residual autocovariance matrices are calculated as in Duchesne (2004). We define the residual cross-covariance matrices between $\{\hat{\mathbf{Z}}_{1t}\}$ and $\{\hat{\mathbf{Z}}_{2t}\}$:

$$\tilde{\mathbf{C}}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k) = \begin{cases} n^{-1}\sum_{t=k+1}^{n}\{\hat{\mathbf{Z}}_{1t} - \mathrm{ave}(\hat{\mathbf{Z}}_1)\}\{\hat{\mathbf{Z}}_{2,t-k} - \mathrm{ave}(\hat{\mathbf{Z}}_2)\}^\top, & k = 0, 1, ..., n-1, \\ \tilde{\mathbf{C}}^\top_{\hat{\mathbf{Z}}_2\hat{\mathbf{Z}}_1}(-k), & k = -n+1, ..., -1, \end{cases}$$

where $\mathrm{ave}(\hat{\mathbf{Z}}_i) = n^{-1}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{it}$, $i = 1, 2$. Under the null hypothesis, $\{\mathbf{Z}_{1t}\}$ and $\{\mathbf{Z}_{2t}\}$ are independent, and it seems natural to study the residual cross-covariances between $\{\hat{\mathbf{Z}}_{1t}\}$ and $\{\hat{\mathbf{Z}}_{2t}\}$ to check causality relations in the second moment.

When the model is correct, using the linearity property of the operator $\mathrm{vech}(\cdot)$, we obtain the relation:

$$n^{-1}\sum_{t=1}^{n}\mathrm{vech}\{\mathbf{L}_{it}^{-1}\boldsymbol{\epsilon}_{it}\boldsymbol{\epsilon}_{it}^\top(\mathbf{L}_{it}^{-1})^\top\} = \mathrm{vech}\left\{n^{-1}\sum_{t=1}^{n}\mathbf{L}_{it}^{-1}\boldsymbol{\epsilon}_{it}\boldsymbol{\epsilon}_{it}^\top(\mathbf{L}_{it}^{-1})^\top\right\},$$

and thus $\mathrm{vech}\{n^{-1}\sum_{t=1}^{n}\boldsymbol{\eta}_{it}\boldsymbol{\eta}_{it}^\top\} \stackrel{\text{a.s.}}{\to} \mathrm{vech}\{\mathbf{I}_{d_i}\}$, given the fact that $\mathrm{var}(\boldsymbol{\eta}_{it}) = \mathbf{I}_{d_i}$, $i = 1, 2$. Thus we replace $\tilde{\mathbf{C}}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)$ by the following dependence measures:

$$\mathbf{C}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k) = \begin{cases} n^{-1}\sum_{t=k+1}^{n}\{\hat{\mathbf{Z}}_{1t} - \mathrm{vech}(\mathbf{I}_{d_1})\}\{\hat{\mathbf{Z}}_{2,t-k} - \mathrm{vech}(\mathbf{I}_{d_2})\}^\top, & k = 0, 1, ..., n-1, \\ \mathbf{C}^\top_{\hat{\mathbf{Z}}_2\hat{\mathbf{Z}}_1}(-k), & k = -n+1, ..., -1. \end{cases}$$

The residual cross-covariance matrices for negative and positive lags are collected in the $\{(2M+1)d_1'd_2'\}\times 1$ vector $\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2} = \left(\mathbf{c}^\top_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(-M), \ldots, \mathbf{c}^\top_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(M)\right)^\top$, where $\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k) = \text{vec}\{\mathbf{C}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)\}$ corresponds to a $(d_1'd_2') \times 1$ random vector.

Another approach, due initially to Ling and Li (1997) for checking multivariate GARCH models, proposes to study the sum of squared (standardized) residuals, $\hat{q}_{it} = \hat{\boldsymbol{\epsilon}}_{it}^\top \hat{\mathbf{V}}_{it}^{-1} \hat{\boldsymbol{\epsilon}}_{it}$, $i = 1, 2$. Given that $\{\hat{q}_{it}, t = 1, \ldots, n\}$ are univariate time series of transformed residuals, cross-covariances can be defined:

$$c_{\hat{q}_1\hat{q}_2}(k) = \begin{cases} n^{-1}\sum_{t=k+1}^{n}\{\hat{q}_{1t} - d_1\}\{\hat{q}_{2,t-k} - d_2\}, & k = 0, 1, ..., n-1, \\ c_{\hat{q}_2\hat{q}_1}(-k), & k = -n+1, ..., -1, \end{cases}$$

and cross-correlations can be computed as:

$$r_{\hat{q}_1\hat{q}_2}(k) = c_{\hat{q}_1\hat{q}_2}(k)\{c_{\hat{q}_1\hat{q}_1}(0)c_{\hat{q}_2\hat{q}_2}(0)\}^{-1/2},$$

with $c_{\hat{q}_1\hat{q}_1}(0) = n^{-1}\sum_{t=1}^{n}(\hat{q}_{1t} - d_1)^2$ and $c_{\hat{q}_2\hat{q}_2}(0) = n^{-1}\sum_{t=1}^{n}(\hat{q}_{2t} - d_2)^2$, noting that $E(q_{it}) = E(\boldsymbol{\epsilon}_{it}^\top \mathbf{V}_{it}^{-1} \boldsymbol{\epsilon}_{it}) = d_i$, $i = 1, 2$. The approach of Ling and Li (1997) reduces the multivariate problem to the computation of cross-covariances or cross-correlations between univariate time series. It is appealing from a computational point of view, but is expected to be less powerful than methods based on multivariate matricial measures (see, e.g., Tse and Tsui (1999) and Tse (2002)). However, as noted in Ling and Li (1997, p. 448), that approach leads to simpler tests in the spirit of portmanteau test statistics in ARMA modelling. See also Li (2004, Section 6.4). Due to its inherent simplicity, we also propose to use that approach for checking causality in variance. We extend the results of Ling and Li (1997) and we establish the properties of the cross-correlations based on transformed residuals. The cross-correlations for negative and positive lags are included in the $(2M+1) \times 1$ random vector $\mathbf{r}_{\hat{q}_1\hat{q}_2} = (r_{\hat{q}_1\hat{q}_2}(-M), \ldots, r_{\hat{q}_1\hat{q}_2}(M))^\top$.

In Section 3, the asymptotic distributions of the dependence measures $\mathbf{C}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)$ and $\mathbf{c}_{\hat{q}_1\hat{q}_2}(k)$, $|k| \leq M$, are established under the null hypothesis of no causality in variance.

3. TEST STATISTICS FOR CAUSALITY IN VARIANCE

3.1 Asymptotic distribution of $n^{1/2}\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}$

Assuming the normality of the white noise process $\{\boldsymbol{\eta}_t\}$, the asymptotic distributions of the residual autocovariances have been studied by Duchesne (2004) in a stationary and ergodic multivariate time series which satisfy equations (3) and (4), for a given i. We now propose to study the residual cross-covariances between two systems, that is between two multivariate time series $\{\mathbf{X}_{1t}\}$ and $\{\mathbf{X}_{2t}\}$. Let $'\otimes'$ be the Kronecker product (see, e.g., Harville (1997)). In order to derive the asymptotic distribution of the residual cross-covariances $n^{1/2}\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}$ under the null hypothesis of no causality in variance, the following lemma is useful.

Lemma 1. *Let $\{\mathbf{X}_{1t}\}$ and $\{\mathbf{X}_{2t}\}$ be two time series generated by equations (3) and (4) such that $E(||\mathbf{Z}_i||^4) < \infty$, where $\mathbf{Z}_{it} = vech\{\mathbf{L}_{it}^{-1}\boldsymbol{\epsilon}_{it}\boldsymbol{\epsilon}_{it}^\top(\mathbf{L}_{it}^{-1})^\top\}$, $i = 1, 2$. If $\{\boldsymbol{\eta}_{1t}\}$ and $\{\boldsymbol{\eta}_{2t}\}$ are independent, then the following asymptotic result holds:*

$$n^{1/2}\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k) \xrightarrow{d} \mathcal{N}_{d_1'd_2'}(\mathbf{0}, \boldsymbol{\Gamma}_{\mathbf{Z}_2\mathbf{Z}_2}(0) \otimes \boldsymbol{\Gamma}_{\mathbf{Z}_1\mathbf{Z}_1}(0)),$$

where $\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k) = \text{vec}\{\mathbf{C}_{\mathbf{Z}_1\mathbf{Z}_2}(k)\}$ denotes the lag-k sample cross-covariance. Furthermore, $\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)$ and $\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k')$ are uncorrelated, $k \neq k'$, and

$$n^{1/2}\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2} \xrightarrow{d} \mathcal{N}_{(2M+1)d_1'd_2'}(\mathbf{0}, \mathbf{I}_{2M+1} \otimes \boldsymbol{\Gamma}_{\mathbf{Z}_2\mathbf{Z}_2}(0) \otimes \boldsymbol{\Gamma}_{\mathbf{Z}_1\mathbf{Z}_1}(0)),$$

where $\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2} = \left(\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}^\top(-M), \ldots, \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}^\top(M)\right)^\top$.

Proof: The proof of Lemma 1 is similar as the one of a lemma given in Duchesne (2004, p. 152). We prove it for positive lags. The proof for negative values of k is similar. Write $\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k) = n^{-1}\sum_{t=k+1}^n \mathbf{w}_{tk}$, $\mathbf{w}_{tk} = \text{vec}(\mathbf{W}_{tk})$, where $\mathbf{W}_{tk} = \mathbf{u}_{1t}\mathbf{u}_{2,t-k}^\top$, with $\mathbf{u}_{it} = \mathbf{Z}_{it} - \text{vech}(\mathbf{I}_{d_i})$, $i = 1, 2$. The asymptotic normality follows using the martingale central limit theorem. Using the independence between $\{\boldsymbol{\eta}_{1t}\}$ and $\{\boldsymbol{\eta}_{2t}\}$, $E(\mathbf{w}_{tk}\mathbf{w}_{tk}^\top) = E\{(\mathbf{u}_{2,t-k} \otimes \mathbf{u}_{1t})(\mathbf{u}_{2,t-k}^\top \otimes \mathbf{u}_{1t}^\top)\} = E\{\mathbf{u}_{2,t-k}\mathbf{u}_{2,t-k}^\top\} \otimes E\{\mathbf{u}_{1t}\mathbf{u}_{1t}^\top\} = \boldsymbol{\Gamma}_{\mathbf{Z}_2\mathbf{Z}_2}(0) \otimes \boldsymbol{\Gamma}_{\mathbf{Z}_1\mathbf{Z}_1}(0)$. It is easy to show that $\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)$ and $\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k')$ are uncorrelated, $k \neq k'$, and to deduce the asymptotic distribution of $n^{1/2}\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}$. $\square$

The result of Lemma 1 holds under a general specification for the distribution of the random vector $\boldsymbol{\eta}_t$. If normality is assumed for each time series, useful simplification occurs. Under normality, $\text{var}\,\text{vec}(\boldsymbol{\eta}_{it}\boldsymbol{\eta}_{it}^\top) = \mathbf{I}_{d_i^2} + \mathbf{K}_{d_id_i}$, $i = 1, 2$, where $\mathbf{K}_{d_id_i}$ denotes the commutation matrix (see, e.g., Harville (1997)). Furthermore, $E(\mathbf{u}_{it}\mathbf{u}_{it}^\top) = 2(\mathbf{G}_{d_i}^\top\mathbf{G}_{d_i})^{-1}$, $i = 1, 2$, where $\mathbf{G}_{d_i}$ represents the duplication matrix (Harville (1997, p. 352)). Thus the asymptotic covariance matrix of $n^{1/2}\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)$ simplifies to the formula $4\{(\mathbf{G}_{d_2}^\top\mathbf{G}_{d_2})^{-1} \otimes (\mathbf{G}_{d_1}^\top\mathbf{G}_{d_1})^{-1}\}$.

Let $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ be estimators of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively, satisfying the condition formulated in relation (5). A standard Taylor's series expansion gives:

$$\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2} = \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2} + \frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}}{\partial \boldsymbol{\theta}^\top}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + \mathbf{o}_P(n^{-1/2}),$$

where $\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}/\partial \boldsymbol{\theta}^\top$ corresponds to a $\{(2M+1)(d_1'd_2') \times (K_1+K_2)\}$ matrix of derivatives, where $K_i = P_i + d_i + d_iP_i$, $i = 1, 2$. For a generic vector $\boldsymbol{\alpha}$, straightforward matrix differentiation yields:

$$
\begin{aligned}
\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\alpha}^\top} &= n^{-1} \sum_{t=k+1}^{n} \frac{\partial \mathrm{vec}(\mathbf{u}_{1t}\mathbf{u}_{2,t-k}^\top)}{\partial \boldsymbol{\alpha}^\top}, \\
&= n^{-1} \sum_{t=k+1}^{n} \left\{ (\mathbf{I}_{d_2'} \otimes \mathbf{u}_{1t}) \frac{\partial \mathbf{u}_{2,t-k}}{\partial \boldsymbol{\alpha}^\top} + (\mathbf{u}_{2,t-k} \otimes \mathbf{I}_{d_1'}) \frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\alpha}^\top} \right\}.
\end{aligned}
$$

See for example Lütkepohl (2005) for useful differentiation rules. The following lemma is needed.

Lemma 2. *Let $\{\mathbf{X}_t\}$ and $\{\mathbf{Y}_t\}$ be two independent stochastic processes, such that $\mathbf{X}_t$ and $\mathbf{Y}_t$ are random matrices of dimension $p \times q$ and $q \times r$, respectively. The process $\{\mathbf{X}_t\}$ is composed of independent random matrices, such that $E(\mathbf{X}_t) = \mathbf{0}$. It is assumed that $E(||\mathbf{X}_t||^2) \leq M_X < \infty$ and $E(||\mathbf{Y}_t||^2) \leq M_Y < \infty$ for certain constants M_X and M_Y. Then $n^{-1}\sum_{t=1}^{n} \mathbf{X}_t\mathbf{Y}_t = \mathbf{O}_P(n^{-1/2})$.*

Proof: Let $\mathbf{l}_i(\mathbf{X}_t)$ be line i of $\mathbf{X}_t$, written as a column vector of dimension $q \times 1$, and $\mathbf{c}_j(\mathbf{Y}_t)$ be column j of $\mathbf{Y}_t$, of dimension $q \times 1$. Write $\mathbf{l}_i^\top(\mathbf{X}_t)\mathbf{c}_j(\mathbf{Y}_t) = \sum_{k=1}^{q} X_{t,ik}Y_{t,kj}$, where $\mathbf{X}_t = (X_{t,ik})_{i=1,\ldots,p;k=1,\ldots,q}$ and $\mathbf{Y}_t = (Y_{t,kj})_{k=1,\ldots,q;j=1,\ldots,r}$. An expansion of $E\left\{n^{-1}\sum_{t=1}^{n}\sum_{k=1}^{q} X_{t,ik}Y_{t,kj}\right\}^2$, the independence assumption between $\{\mathbf{X}_t\}$ and $\{\mathbf{Y}_t\}$, and Proposition 6.2.3 of Brockwell and Davis (1991) prove the lemma. □

Under the hypothesis of independence between $\{\boldsymbol{\eta}_{1t}\}$ and $\{\boldsymbol{\eta}_{2t}\}$, it follows that the stochastic processes $\{\mathbf{u}_{1t}\}$ and $\{\partial \mathbf{u}_{2,t-k}/\partial \boldsymbol{\alpha}^\top\}$ are independent. Similarly, we note the stochastic independence between $\{\mathbf{u}_{2,t-k}\}$ and $\{\partial \mathbf{u}_{1t}/\partial \boldsymbol{\alpha}^\top\}$. Using the assumption on the innovation processes $\{\boldsymbol{\eta}_{1t}\}$ and $\{\boldsymbol{\eta}_{2t}\}$, and using Lemma 2, it follows that:

$$\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\alpha}^\top} = \mathbf{O}_P(n^{-1/2}).$$

Thus:

$$\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2} = \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2} + \mathbf{o}_P(n^{-1/2}).$$

Consequently, $n^{1/2}\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}$ and $n^{1/2}\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}$ converge toward the same asymptotic distribution under the null hypothesis. This leads to the following proposition.

Proposition 1. *Let $\{\mathbf{X}_{1t}\}$ and $\{\mathbf{X}_{2t}\}$ be two stationary processes which satisfy equations (3) and (4). Suppose that the corresponding innovation processes satisfy $E(||\mathbf{u}_{it}||^4) < \infty$, $\mathbf{u}_{it} = vech(\boldsymbol{\eta}_{it}\boldsymbol{\eta}_{it}^\top - \mathbf{I}_{d_i})$, $i = 1, 2$. Assume that the processes are fitted by estimating methods satisfying the condition given by relation (5). If the two stochastic processes $\{\boldsymbol{\eta}_{1t}\}$ and $\{\boldsymbol{\eta}_{2t}\}$ are independent, then $n^{1/2}\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}$ and $n^{1/2}\mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}$ asymptotically follow the same distribution.*

The results of Proposition 1 are different from the ones on diagnostic checking, since the asymptotic variance-covariance matrix of the residual autocovariances involves the expectations of their derivatives, which are bounded in probability under certain conditions. As a result of the independence between $\{\boldsymbol{\eta}_{1t}\}$ and $\{\boldsymbol{\eta}_{2t}\}$, the derivatives in the present framework are $\mathbf{O}_P(n^{-1/2})$. Proposition 1 generalizes in a multivariate framework a theorem of Cheung and Ng (1996, Theorem 1).

Similar arguments allow us to derive the asymptotic distribution of $\mathbf{r}_{\hat{q}_1\hat{q}_2}$. In fact, it is easy to show that $n^{1/2}\mathbf{r}_{q_1q_2}$ converges in distribution toward a $\mathcal{N}_{2M+1}(\mathbf{0}, \mathbf{I}_{2M+1})$. In the following expansion:

$$\mathbf{r}_{\hat{q}_1\hat{q}_2} = \mathbf{r}_{q_1q_2} + \frac{\partial \mathbf{r}_{q_1q_2}}{\partial \boldsymbol{\alpha}^\top}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + \mathbf{o}_P(n^{-1/2}),$$

we can show that $\partial \mathbf{r}_{q_1q_2}/\partial \boldsymbol{\alpha}^\top = \mathbf{O}_P(n^{-1/2})$ and thus:

$$\mathbf{r}_{\hat{q}_1\hat{q}_2} = \mathbf{r}_{q_1q_2} + \mathbf{o}_P(n^{-1/2}).$$

It follows that the asymptotic distributions of $n^{1/2}\mathbf{r}_{\hat{q}_1\hat{q}_2}$ and $n^{1/2}\mathbf{r}_{q_1q_2}$ are the same. This is formally presented in the next proposition.

Proposition 2. *Let $\{\mathbf{X}_{1t}\}$ and $\{\mathbf{X}_{2t}\}$ be two stationary processes which satisfy equations (3) and (4). Suppose that the corresponding innovation processes satisfy $E(||\mathbf{u}_{it}||^4) < \infty$, $\mathbf{u}_{it} = vech(\boldsymbol{\eta}_{it}\boldsymbol{\eta}_{it}^\top - \mathbf{I}_{d_i})$, $i = 1, 2$. Assume that the processes are fitted by estimating methods satisfying the condition given by relation (5). If the two stochastic processes $\{\boldsymbol{\eta}_{1t}\}$ and $\{\boldsymbol{\eta}_{2t}\}$ are independent, then $n^{1/2}\mathbf{r}_{\hat{q}_1\hat{q}_2}$ and $n^{1/2}\mathbf{r}_{q_1q_2}$ asymptotically follow the same distribution.*

The proof of Proposition 2 is similar as the one of Proposition 1, and analogous arguments to those of Ling and Li (1997) are used in order to manipulate the transformed residuals $\hat{q}_i$, $i = 1, 2$. As in Proposition 1, the key part is the independence assumption between the processes $\{\boldsymbol{\eta}_{1t}\}$ and $\{\boldsymbol{\eta}_{2t}\}$. Proposition 2 complements results of Ling and Li (1997, Section 3). See also Li (2004, p. 125).

3.2 Test statistics at individual lags

For a finite set of lags $\{k_1, \ldots, k_m\}$, the asymptotic distributions of the random vectors $n^{1/2}\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}$ and $n^{1/2}\mathbf{r}_{\hat{q}_1\hat{q}_2}$ given by Propositions 1 and 2 are particularly simple to use in the construction of test statistics for the null hypothesis of no causality in variance. Since $n^{1/2}\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k_j)$, $j = 1, \ldots, m$ are asymptotically independent and identically distributed $\mathcal{N}_{d_1'd_2'}(\mathbf{0}, \boldsymbol{\Gamma}_{\mathbf{Z}_2\mathbf{Z}_2}(0) \otimes \boldsymbol{\Gamma}_{\mathbf{Z}_1\mathbf{Z}_1}(0))$ under the null hypothesis, it is natural to construct the test statistic:

$$\begin{aligned}\mathrm{ER}_n(k) &= n\mathbf{c}^{\top}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)\{\mathbf{C}^{-1}_{\hat{\mathbf{Z}}_2\hat{\mathbf{Z}}_2}(0)\otimes\mathbf{C}^{-1}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_1}(0)\}\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k),\\ &= n\mathrm{tr}\{\mathbf{C}^{\top}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)\mathbf{C}^{-1}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_1}(0)\mathbf{C}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)\mathbf{C}^{-1}_{\hat{\mathbf{Z}}_2\hat{\mathbf{Z}}_2}(0)\}.\end{aligned} \tag{6}$$

A similar test statistic for testing causality in mean has been considered in El Himdi and Roy (1997). The second equality in the preceding formulas follows invoking Harville (1997, Theorem 16.2.2). Using the trace representation is often more computationally efficient than the one using the Kronecker product. Under the null hypothesis the test statistic $\mathrm{ER}_n(k)$ converges toward a $\chi^2_{d'_1 d'_2}$ distribution. The null hypothesis is rejected for large values of the test statistic, that is if $\mathrm{ER}_n(k) > \chi^2_{d'_1 d'_2,1-\alpha}$, where $\chi^2_{d,1-\alpha}$ denotes the quantile of order $1-\alpha$ of the χ^2_d random variable.

Similarly, Proposition 2 suggests to construct the test statistic:

$$\mathrm{LL}_n(k) = nr^2_{\hat{q}_1\hat{q}_2}(k), \tag{7}$$

which converges in distribution toward a χ^2_1 random variable under the null hypothesis. The critical region is given by $\{\mathrm{LL}_n(k) > \chi^2_{1,1-\alpha}\}$.

As in the univariate case for testing causality in variance, or for testing causality in mean between multivariate time series, the test statistics $\mathrm{ER}_n(k)$ and $\mathrm{LL}_n(k)$ use the factor n^{-1} as the asymptotic variance. However, the exact variance can be smaller. A better approximation is provided by the factor $(n-|k|)n^{-2}$. Consequently we introduce the modified test statistics:

$$\mathrm{ER}^*_n(k) = \omega_k \mathrm{ER}_n(k), \tag{8}$$

$$\mathrm{LL}^*_n(k) = \omega_k \mathrm{LL}_n(k). \tag{9}$$

The factors $\omega^{(1)}_k = n/(n-|k|)$ or $\omega^{(1)}_k = (n+2)/(n-|k|)$ may be recommended (see, e.g., Haugh (1976) and Mcleod and Li (1983), among others). Note that the modified versions are strictly larger than the original test statistics, that is $\mathrm{ER}^*_n(k) > \mathrm{ER}_n(k)$ and $\mathrm{LL}^*_n(k) > \mathrm{LL}_n(k)$, but these modified test statistics are asymptotically equivalent to the original ones.

In order to consider all the lags in the set $\{k_1, \ldots, k_m\}$, global test statistics may rely on $\mathrm{ER}^*_n(k)$ or $\mathrm{LL}^*_n(k)$, $k \in \{k_1, \ldots, k_m\}$. According to Propositions 1 and 2, these test statistics are asymptotically independent. Thus in that situation it is particularly simple to modify the significance level in order to obtain a global significance level α. In fact, for each approach, it suffices to fix the marginal nominal level of each test statistic to the value $\alpha_0 = 1-(1-\alpha)^{1/m}$. For example, for the approach based on $\mathrm{ER}^*_n(k)$, the null hypothesis is rejected if $\mathrm{ER}_n^*(k_j) > \chi^2_{d_1' d_2', 1-\alpha_0}$ for at least one k_j. Similarly, if one adopts the approach based on $\mathrm{LL}_n^*(k)$, the procedure rejects the null hypothesis if $\mathrm{LL}_n^*(k_j) > \chi^2_{1,1-\alpha_0}$ for at least one k_j.

3.3 Procedures for detecting the directions of causality in variance

In addition to the test statistics at individual lags $\mathrm{ER}_n(k)$, $\mathrm{LL}_n(k)$ and their modified versions, test procedures for more specific causality hypotheses can be obtained easily. Based on the first approach, we introduce the test statistics:

$$\mathrm{ER}_n(L:U) = n \sum_{k=L}^{U} \mathrm{tr}\{\mathbf{C}^\top_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)\mathbf{C}^{-1}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_1}(0)\mathbf{C}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)\mathbf{C}^{-1}_{\hat{\mathbf{Z}}_2\hat{\mathbf{Z}}_2}(0)\}, \quad (10)$$

$$\mathrm{ER}^*_n(L:U) = n \sum_{k=L}^{U} \omega_k \mathrm{tr}\{\mathbf{C}^\top_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)\mathbf{C}^{-1}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_1}(0)\mathbf{C}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}(k)\mathbf{C}^{-1}_{\hat{\mathbf{Z}}_2\hat{\mathbf{Z}}_2}(0)\}, \quad (11)$$

with $L \leq U < n$. Using the transformed residuals, we also introduce:

$$\mathrm{LL}_n(L:U) = n \sum_{k=L}^{U} r^2_{\hat{q}_1\hat{q}_2}(k), \quad (12)$$

$$\mathrm{LL}^*_n(L:U) = n \sum_{k=L}^{U} \omega_k r^2_{\hat{q}_1\hat{q}_2}(k). \quad (13)$$

The test statistics $\mathrm{ER}_n(L:U)$ and $\mathrm{ER}^*_n(L:U)$ converge toward $\chi^2_{(U-L+1)d_1'd_2'}$ random variables under the null hypothesis, while the asymptotic null distribution of the test statistics $\mathrm{LL}_n(L:U)$ and $\mathrm{LL}^*_n(L:U)$ are χ^2_{U-L+1}. The test statistics (10) and (12) reduce to the test statistic (1) of Cheung and Ng (1996) when the two time series are univariate. Note that these authors also investigated modified versions.

We now discuss possible choices of L and U. Suppose that we want to test the null hypothesis of non causality in variance against causality in variance from $\{\mathbf{X}_{1t}\}$ to $\{\mathbf{X}_{2t}\}$. Then the choice $(L, U) = (-M, -1)$, $M < n$ is appropriate. On the other hand, $(L, U) = (1, M)$ can be used when the alternative specifies causality in variance from $\{\mathbf{X}_{2t}\}$ to $\{\mathbf{X}_{1t}\}$. When no precise direction of causality in variance is known a priori, generalizations of the global portmanteau test statistics of Cheung and Ng (1996) can be proposed. They correspond to the choice $(L,U)=(-M,M)$ and they are noted:

$$\begin{aligned} \mathrm{ER}_{n,M} &= \mathrm{ER}_n(-M:M), & (14)\\ \mathrm{ER}^*_{n,M} &= \mathrm{ER}^*_n(-M:M), & (15)\\ \mathrm{LL}_{n,M} &= \mathrm{LL}_n(-M:M), & (16)\\ \mathrm{LL}^*_{n,M} &= \mathrm{LL}^*_n(-M:M). & (17)\end{aligned}$$

In practice, it is expected that the test statistics $\mathrm{LL}_{n,M}$ and $\mathrm{LL}_n(k)$ and their modified versions should be particularly convenient since statistical software such as R or S-PLUS already include built-in functions to compute the residual cross-correlations. The analyst just need to retrieve the residuals from the fitted models, to transform them, and finally to calculate lag-k cross-correlations. However, they are expected to be less powerful than the test statistics $\mathrm{ER}_{n,M}$, $\mathrm{ER}_n(k)$ and their modified versions. In the next section, the test statistics proposed in this section are studied empirically. In particular, the power issues are investigated.

4. SIMULATION EXPERIMENTS

In the previous section, we have introduced two classes of test statistics which should prove useful for testing causality in variance. From a practical point of view, it appears useful to study their finite sample properties, in particular their exact levels. The power comparisons between the global test statistics $\mathrm{LL}_{n,M}$ and $\mathrm{ER}_{n,M}$, and also between the one-lag test statistics $\mathrm{LL}_n(k)$ and $\mathrm{ER}_n(k)$ are also of special interest. These issues are studied in this section.

4.1 Description of the experiments

For given bivariate time series models described below, we examined the empirical frequencies of rejection of the null hypothesis when the latter is true by test procedures with two different nominal levels (5 and 10 percent), for each of two series lengths ($n = 1000, 2000$). We included in our study the test statistics at individual lags $\mathrm{ER}_n(k)$, $\mathrm{ER}^*_n(k)$, $\mathrm{LL}_n(k)$ and $\mathrm{LL}^*_n(k)$, for $k=-10,\ldots,-1,0,1,\ldots,10$. The portmanteau test procedures $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$ were also considered in the simulation experiments for $M=1,\ldots,10$.

To investigate the finite sample properties of the test statistics under the null hypothesis, we simulated the following data generating processes (DGPs):

$$\begin{aligned} \mathbf{X}_{it} &= \mathbf{L}_{it}\boldsymbol{\eta}_{it}, & (18)\\ \mathbf{V}_{it}=\mathbf{L}_{it}\mathbf{L}_{it}^\top &= \boldsymbol{\Theta}+\mathbf{A}_i^\top\left(\boldsymbol{\epsilon}_{i,t-1}\boldsymbol{\epsilon}_{i,t-1}^\top\right)\mathbf{A}_i+\mathbf{B}_i^\top\mathbf{V}_{i,t-1}\mathbf{B}_i,\ i=1,2. & (19)\end{aligned}$$

We considered two bivariate stochastic processes, that is $d_1 = d_2 = 2$. The conditional variance specification (19) corresponds to a BEKK(1,1) model. The stochastic processes $\{\boldsymbol{\eta}_{it}\}$, $i = 1, 2$, were assumed to be independent Gaussian white noises $\mathcal{N}_2(\mathbf{0}, \mathbf{I}_2)$. We investigated three situations in the level study. Consider the model parameters defined in Table 1. For each case, the model parameters are given as follows:

$$
\begin{aligned}
L_1 :& \quad \mathbf{A}_1 = \mathbf{A}_2 = \boldsymbol{\Phi}_1,\ \mathbf{B}_1 = \mathbf{B}_2 = \boldsymbol{\Phi}_2,\\
L_2 :& \quad \mathbf{A}_1 = \boldsymbol{\Phi}_1,\ \mathbf{A}_2 = \boldsymbol{\Phi}_4,\ \mathbf{B}_1 = \boldsymbol{\Phi}_2,\ \mathbf{B}_2 = \boldsymbol{\Phi}_4,\\
L_3 :& \quad \mathbf{A}_1 = \mathbf{A}_2 = \boldsymbol{\Phi}_1,\ \mathbf{B}_1 = \boldsymbol{\Phi}_2,\ \mathbf{B}_2 = \boldsymbol{\Phi}_3.
\end{aligned}
$$

Similar models for the conditional variances have been considered in Duchesne and Lalancette (2002, 2003). Under the DGP denoted L_1, a diagonal BEKK GARCH(1,1) is simulated. The coefficients are the same for each time series. Under L_2, a diagonal BEKK is considered for $\{\mathbf{X}_{1t}\}$ but a non-diagonal BEKK GARCH(1,1) is formulated for $\{\mathbf{X}_{2t}\}$. Under the third model L_3, a similar model to L_1 is used but the volatility is more persistent for $\{\mathbf{X}_{2t}\}$ than for $\{\mathbf{X}_{1t}\}$.

In order to study the power properties of the test statistics, the conditional variances in the model (18) and (19) are assumed to be:

$$
\begin{aligned}
\mathbf{V}_{1t} &= \boldsymbol{\Theta} + \mathbf{A}_1^\top \left(\boldsymbol{\epsilon}_{1,t-1}\boldsymbol{\epsilon}_{1,t-1}^\top\right) \mathbf{A}_1 + \mathbf{B}_1^\top \mathbf{V}_{1,t-1}\mathbf{B}_1 + \\
&\quad \delta_{12}\mathbf{C}_1^\top \left(\boldsymbol{\epsilon}_{2,t-l_1}\boldsymbol{\epsilon}_{2,t-l_1}^\top\right) \mathbf{C}_1 + \gamma_{12}\mathbf{D}_1^\top \mathbf{V}_{2,t-l_1}\mathbf{D}_1, \qquad (20)\\
\mathbf{V}_{2t} &= \boldsymbol{\Theta} + \mathbf{A}_2^\top \boldsymbol{\epsilon}_{2,t-1}\boldsymbol{\epsilon}_{2,t-1}^\top \mathbf{A}_2 + \mathbf{B}_2^\top \mathbf{V}_{2,t-1}\mathbf{B}_2 + \\
&\quad \delta_{21}\mathbf{C}_2^\top \left(\boldsymbol{\epsilon}_{1,t-l_2}\boldsymbol{\epsilon}_{2,t-l_2}^\top\right) \mathbf{C}_2 + \gamma_{21}\mathbf{D}_2^\top \mathbf{V}_{1,t-l_2}\mathbf{D}_2. \qquad (21)
\end{aligned}
$$

When $\delta_{12} \neq 0$ or $\gamma_{12} \neq 0$, volatility spillover exists from $\{\mathbf{X}_{2t}\}$ to $\{\mathbf{X}_{1t}\}$, and l_1 controls the lag delay in the volatility dynamics. Similar interpretations are given for the other parameters δ_{21}, γ_{21} and l_2.

We investigated four models for the alternatives, noted A_i, $i \in \{1, \ldots, 4\}$. The parameters $\mathbf{A}_i$, $\mathbf{B}_i$, $i = 1, 2$, for alternatives A_i, $i = 1, 2, 3$, were defined as in model L_1. The other

Table 1. Model parameters in the level and power studies

$$
\boldsymbol{\Theta} = \begin{pmatrix} 0.3 & 0.1 \\ 0.1 & 0.3 \end{pmatrix}, \quad \boldsymbol{\Phi}_1 = \begin{pmatrix} 0.5 & 0.0 \\ 0.0 & 0.5 \end{pmatrix}, \quad \boldsymbol{\Phi}_2 = \begin{pmatrix} 0.4 & 0.0 \\ 0.0 & 0.4 \end{pmatrix},
$$
$$
\boldsymbol{\Phi}_3 = \begin{pmatrix} 0.8 & 0.0 \\ 0.0 & 0.8 \end{pmatrix}, \quad \boldsymbol{\Phi}_4 = \begin{pmatrix} 0.4 & 0.1 \\ 0.1 & 0.4 \end{pmatrix},
$$

parameters were given by $\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{\Phi}_1$, $\mathbf{D}_1 = \mathbf{D}_2 = \mathbf{\Phi}_3$ and $l_1 = l_2 = 1$. The parameters δ_{ij}, γ_{ij}, $i, j \in \{1, 2\}$, were specified as follows:

$$
\begin{aligned}
A_1: &\quad \delta_{12} = \gamma_{12} = \delta_{21} = \gamma_{21} = 0.17, \\
A_2: &\quad \delta_{12} = \gamma_{12} = 0.25,\ \delta_{21} = \gamma_{21} = 0.0, \\
A_3: &\quad \delta_{12} = \gamma_{12} = 0.11,\ \delta_{21} = \gamma_{21} = 0.17.
\end{aligned}
$$

Under the alternative A_4, the model parameters for $\mathbf{A}_i$, $\mathbf{B}_i$, $i = 1, 2$ were chosen as in the DGP denoted L_2. The other model parameters were:

$$A_4: \quad \mathbf{C}_1 = \mathbf{C}_2 = \mathbf{\Phi}_1,\ \mathbf{D}_1 = \mathbf{D}_2 = \mathbf{\Phi}_3,\ l_1 = l_2 = 4,\ \delta_{12} = \gamma_{12} = \delta_{21} = \gamma_{21} = 0.17.$$

Under the alternative A_1, there exists causality in variance from $\{\mathbf{X}_{2t}\}$ to $\{\mathbf{X}_{1t}\}$ and from $\{\mathbf{X}_{1t}\}$ to $\{\mathbf{X}_{2t}\}$, with respect to $\mathcal{F}_{t-1}$. Under alternative A_2, there exists unilateral volatility spillover from $\{\mathbf{X}_{2t}\}$ to $\{\mathbf{X}_{1t}\}$ (but not from $\{\mathbf{X}_{1t}\}$ to $\{\mathbf{X}_{2t}\}$). Alternative A_3 is similar to alternative A_1, but the coefficients δ_{12} and γ_{12} are smaller than δ_{21} and γ_{21}, creating different causality in variance dynamics. Finally, under alternative A_4, causality in variance in both directions occurs, but with a four period lag delay. All the models have been estimated using multivariate GARCH models written in BEKK(1,1) form, using the S-PLUS function `mgarch` in the `Finmetrics` module. In order to obtain the maximum likelihood estimators, the BHHH algorithm of Berndt *et al.* (1974) has been used, specifying a Gaussian likelihood. That is similar to the simulation experiments performed by Hong (2001) in the univariate case. A total of 1000 independent realizations were generated for each experiment. The standard errors of the empirical levels based on 1000 independent realizations are 0.689% and 0.949% for the nominal levels 5% and 10%, respectively.

4.2 Discussion of the experiments under the null hypothesis

We first examine the performance for the levels of the various test statistics. Tables 2-4 show the rejection rates for the global test statistics with maximal lag order M, $M = 1, \ldots, 10$. Under DGP L_1, the empirical levels were generally very close to the nominal levels for every value of M at the 5% nominal level and $n = 1000$, except those of $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$ for small values of M, that is $M \leq 2$. A small improvement has been observed for the modified portmanteau test statistics. At the 10% nominal level, the rejection rates of the test statistics $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$ were under the nominal levels when $n = 1000$, but for $M \geq 3$ the empirical levels were close to the 10% level. When $n = 2000$, the rejection rates were generally satisfying, except maybe when $M = 1$ for the test statistics $\mathrm{ER}_{n,M}$ and $\mathrm{ER}^*_{n,M}$, and a small deterioration has been observed for $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$ when $M \geq 9$. Under the DGPs L_2 and L_3, similar conclusions have been reached.

TABLE 2. Empirical levels, model L_1, nominal levels 5% and 10%, $n = 1000, 2000$, for the global test statistics $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
M	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
1	5.4	5.4	3.6	3.6	4.7	4.7	5.1	5.1	10.1	10.1	8.2	8.2	8.5	8.5	9.7	9.7
2	5.4	5.5	4.2	4.2	5.0	5.0	5.7	5.7	10.0	10.0	7.4	7.4	9.0	9.0	10.6	10.6
3	5.7	5.8	4.9	4.9	4.8	4.8	4.5	4.5	10.4	10.4	9.5	9.5	9.5	9.6	10.4	10.4
4	5.3	5.4	5.0	5.0	5.3	5.3	6.3	6.3	10.7	10.9	9.5	9.5	10.5	10.5	10.8	10.9
5	5.5	5.9	4.6	4.6	5.8	5.8	5.3	5.5	9.3	9.7	9.7	10.0	10.2	10.4	10.3	10.4
6	4.7	5.3	4.9	5.1	5.5	5.5	5.0	5.0	10.0	10.5	9.7	10.0	10.6	11.0	9.9	10.2
7	5.3	5.7	5.0	5.2	4.9	5.0	4.2	4.3	10.2	10.8	9.6	9.6	9.3	9.5	9.7	10.1
8	5.6	5.7	4.5	4.6	4.4	4.6	4.5	4.5	9.7	10.6	8.8	9.1	9.4	9.4	9.5	9.6
9	5.3	6.0	5.3	5.6	5.1	5.4	4.8	4.9	9.7	10.6	9.4	9.7	8.8	9.1	8.6	8.8
10	5.4	5.8	5.9	6.1	5.7	6.0	4.4	4.7	9.9	11.1	10.2	10.3	9.8	10.0	8.8	8.8

TABLE 3. Empirical levels, model L_2, nominal levels 5% and 10%, $n = 1000, 2000$, for the global test statistics $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
M	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
1	5.6	5.6	3.9	3.9	4.5	4.5	5.3	5.3	10.3	10.3	8.2	8.2	8.4	8.4	9.6	9.6
2	5.5	5.6	3.9	3.9	5.0	5.0	5.5	5.5	9.3	9.5	7.7	7.8	9.3	9.4	10.3	10.3
3	5.8	5.8	4.7	4.7	4.9	4.9	4.4	4.4	9.9	10.1	9.4	9.5	9.5	9.6	10.3	10.3
4	5.2	5.2	4.8	4.9	5.4	5.4	6.4	6.4	10.2	10.5	9.2	9.4	10.3	10.3	10.6	10.6
5	5.9	6.2	4.6	4.7	5.6	5.8	5.7	5.8	9.2	9.4	9.8	9.8	10.0	10.3	10.5	10.5
6	5.0	5.1	5.3	5.3	5.5	5.7	5.2	5.2	10.3	10.8	9.8	9.9	10.4	10.6	10.2	10.2
7	5.0	5.2	5.3	5.4	4.7	5.1	4.3	4.5	10.2	10.6	9.1	9.5	9.1	9.4	10.3	10.4
8	5.8	6.0	4.3	4.5	4.2	4.7	4.6	4.6	10.1	10.6	9.0	9.4	9.5	9.6	9.3	9.5
9	5.4	6.0	4.7	5.1	4.9	5.3	4.7	4.9	10.2	10.6	9.3	9.6	9.1	9.6	8.9	8.9
10	5.6	5.7	5.7	5.8	5.7	5.8	4.4	4.6	10.0	11.2	10.1	10.5	9.6	10.0	8.9	9.0

TABLE 4. Empirical levels, model L_3, nominal levels 5% and 10%, $n = 1000, 2000$, for the global test statistics $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$, $M = 1, \dots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
M	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
1	5.6	5.6	4.0	4.0	3.8	3.8	4.7	4.7	9.8	9.8	8.1	8.1	9.5	9.5	9.6	9.6
2	5.5	5.5	3.6	3.6	4.5	4.5	5.7	5.7	9.7	9.8	7.7	7.7	8.4	8.4	10.5	10.5
3	6.3	6.3	5.2	5.3	4.9	5.0	4.6	4.7	10.0	10.2	9.5	9.7	9.1	9.1	9.4	9.5
4	5.3	5.3	5.1	5.1	5.3	5.5	5.8	5.9	10.9	11.0	9.6	9.6	10.0	10.2	10.5	10.5
5	6.2	6.2	5.0	5.1	6.0	6.0	5.5	5.5	9.8	10.2	9.9	9.9	9.7	10.1	10.1	10.2
6	5.3	5.5	5.4	5.5	5.4	5.5	5.0	5.1	10.1	10.6	9.8	10.1	10.3	10.4	9.6	9.6
7	5.0	5.3	5.0	5.1	5.1	5.1	4.4	4.4	10.7	11.0	9.7	9.8	9.5	9.5	9.5	9.6
8	5.6	5.8	4.6	4.6	4.9	5.3	4.4	4.5	9.8	10.8	9.7	9.8	9.4	9.7	9.0	9.2
9	5.3	5.8	5.1	5.1	5.2	5.4	4.4	4.4	9.8	10.8	9.9	10.2	8.8	9.0	8.4	8.8
10	5.1	5.4	5.3	5.5	5.5	5.7	4.4	4.4	9.7	11.3	10.2	10.6	9.7	10.1	9.7	8.7

Tables 5-7 present the rejection rates for the test statistics at individual lags. In general, the χ^2 approximation provides a good approximation for the test statistics $\mathrm{ER}_n(k)$ and $\mathrm{ER}^*_n(k)$. Similar behaviors are observed for $\mathrm{LL}_n(k)$ and $\mathrm{LL}^*_n(k)$, particularly when $n = 2000$. That conclusion is different than the one of El Himdi and Roy (1997), who observed significant improvements with the modified test statistics over the original ones. That may be due to the larger sample sizes in our simulation experiments. When testing for causality in mean, El Himdi and Roy (1997) investigated studies with sample sizes as low as $n = 50$ observations. Such sample sizes are small in modelling multivariate GARCH models. However, adjusting multivariate GARCH models to economic and financial time series data, it is realistic to have sample sizes as large as the ones considered in our simulation experiments. Consequently, to use the modified versions may be unnecessary in our present framework.

4.3 Discussion of the experiments under the alternative hypothesis

We now turn to the power properties. Since the empirical levels are close to the nominal levels, all the empirical powers have been calculated with the asymptotic critical values. Tables 8-11 present the empirical powers for the global test statistics. In general, we observe that the empirical powers of the original test statistics and their corresponding modified versions are largely similar. With a large value of M, a slight advantage has been observed for

Table 5. Empirical levels, model L_1, nominal levels 5% and 10%, $n = 1000, 2000$, for the test statistics at individual lags $\mathrm{ER}_n(k)$, $\mathrm{ER}^*_n(k)$, $\mathrm{LL}_n(k)$ and $\mathrm{LL}^*_n(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
k	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
-10	6.6	6.8	6.1	6.2	5.9	6.0	5.3	5.4	10.6	11.1	10.2	10.6	10.6	11.1	10.6	10.8
-9	5.1	5.3	4.9	4.9	4.5	4.8	4.2	4.2	9.5	9.9	8.7	9.1	10.3	10.6	8.9	9.0
-8	5.3	5.5	5.8	5.9	5.3	5.4	4.5	4.6	10.5	11.0	10.8	10.9	8.8	8.9	9.4	9.4
-7	4.6	4.9	4.8	5.2	5.9	6.2	5.4	5.4	10.1	10.5	8.9	9.0	10.9	11.1	11.1	11.1
-6	5.8	5.8	4.2	4.4	4.6	4.6	5.1	5.3	10.6	11.0	10.0	10.1	8.8	8.9	9.4	9.5
-5	3.9	4.1	5.1	5.2	4.5	4.5	5.0	5.0	8.8	8.9	9.1	9.1	8.4	8.4	10.8	10.8
-4	5.7	5.8	4.5	4.5	5.4	5.8	4.3	4.3	10.4	10.5	9.9	10.0	10.2	10.3	8.7	8.7
-3	6.2	6.2	5.7	5.7	4.1	4.2	6.0	6.0	10.2	10.3	11.8	11.9	7.8	7.8	11.7	11.7
-2	5.6	5.7	5.3	5.3	5.6	5.6	5.2	5.2	10.3	10.4	9.3	9.3	9.3	9.4	11.3	11.4
-1	5.8	5.9	4.7	4.8	4.6	4.6	5.4	5.5	10.4	10.5	9.8	9.9	9.5	9.7	10.9	10.9
0	6.0	6.0	3.5	3.5	4.9	4.9	5.4	5.4	9.5	9.5	8.4	8.4	9.7	9.7	9.7	9.7
1	6.6	6.7	5.3	5.3	5.7	5.7	4.9	4.9	11.9	12.0	10.7	10.8	10.1	10.1	10.8	10.8
2	4.7	4.7	3.3	3.3	5.5	5.5	5.2	5.2	9.6	9.6	7.7	7.7	10.6	10.6	10.5	10.5
3	5.7	5.7	6.6	6.6	5.2	5.2	5.0	5.0	10.7	10.7	11.0	11.0	10.7	10.7	10.3	10.3
4	6.7	6.7	5.3	5.5	5.7	5.7	6.9	6.9	10.8	11.1	9.9	10.0	11.5	11.8	11.7	11.7
5	5.3	5.6	3.6	3.8	5.8	5.9	4.2	4.2	10.1	10.1	8.8	9.1	10.0	10.0	10.4	10.4
6	5.5	5.6	5.1	5.1	4.0	4.1	4.1	4.1	10.7	11.1	10.6	10.6	9.3	9.4	8.1	8.1
7	6.0	6.1	4.0	4.0	4.3	4.3	5.1	5.1	10.8	11.3	8.2	8.3	8.0	8.4	9.9	9.9
8	6.1	6.2	5.0	5.1	4.8	4.9	4.6	4.7	10.5	10.7	9.8	9.9	10.8	10.9	8.2	8.2
9	5.2	5.5	5.6	5.8	4.7	4.7	5.2	5.2	10.1	10.2	11.0	11.0	10.0	10.2	9.3	9.4
10	6.2	6.3	3.2	3.3	5.5	5.7	4.6	4.6	10.9	10.6	7.4	7.7	11.1	11.2	8.8	9.0

the modified versions when $n = 1000$, but for $n = 2000$ the empirical powers are essentially the same; that was expected, since the modified test statistics are asymptotically equivalent to the unmodified ones. In general, the test statistics $\mathrm{ER}_{n,N}$ and $\mathrm{ER}^*_{n,N}$ were more powerful than the Ling and Li test statistics $\mathrm{LL}_{n,N}$ and $\mathrm{LL}^*_{n,N}$, at least for the chosen alternatives.

We now discuss the simulation results in more details. Under the alternative A_1, causality in variance exists in both directions. As a function of M, the test statistics have a rapidly increasing power with maximal powers reached at $M = 2$ or $M = 3$. The test statistics $\mathrm{ER}_{n,M}$ and $\mathrm{ER}^*_{n,M}$ were always more powerful than the Ling and Li test statistics $\mathrm{LL}_{n,M}$

TABLE 6. Empirical levels, model L_2, nominal levels 5% and 10%, $n = 1000, 2000$, for the test statistics at individual lags $\mathrm{ER}_n(k)$, $\mathrm{ER}^*_n(k)$, $\mathrm{LL}_n(k)$ and $\mathrm{LL}^*_n(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
k	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
-10	6.3	6.4	5.4	5.4	5.9	6.0	5.4	5.5	10.3	11.0	10.7	11.0	11.0	11.1	11.0	11.1
-9	4.8	5.1	4.8	5.1	4.3	4.5	4.4	4.4	10.0	10.2	8.7	9.1	10.3	10.6	8.9	9.0
-8	5.1	5.2	5.7	6.0	5.6	5.7	4.5	4.6	10.2	11.1	10.7	10.7	8.6	8.8	8.9	9.0
-7	5.1	5.2	4.3	4.5	6.0	6.1	5.4	5.4	10.6	10.8	9.3	9.4	10.9	11.2	10.9	11.1
-6	6.0	6.1	4.5	4.5	4.6	4.6	5.0	5.0	10.6	10.9	10.1	10.1	8.6	8.7	9.5	9.6
-5	4.3	4.4	5.0	5.0	4.1	4.3	4.8	4.8	9.3	9.5	9.1	9.1	8.5	8.6	10.7	10.7
-4	5.7	5.8	4.2	4.2	6.0	6.1	4.0	4.1	10.0	10.2	9.9	10.1	10.2	10.3	8.8	8.8
-3	6.1	6.3	5.6	5.6	4.3	4.3	6.0	6.0	9.9	10.1	11.4	11.6	7.8	7.9	11.3	11.3
-2	7.3	5.3	5.3	5.3	5.7	5.8	5.6	5.6	10.2	10.3	9.3	9.4	9.4	9.4	11.6	11.6
-1	5.9	5.9	4.8	4.8	4.5	4.5	5.5	5.5	11.2	11.2	10.6	10.6	9.7	9.7	10.8	10.8
0	6.3	6.3	3.4	3.4	4.8	4.8	5.4	5.4	9.7	9.7	8.2	8.2	9.7	9.7	9.6	9.6
1	7.1	7.1	5.0	5.0	5.5	5.5	5.0	5.0	11.8	11.9	10.2	10.2	9.9	9.9	10.7	10.7
2	4.3	4.4	3.3	3.3	5.4	5.5	5.2	5.2	9.4	9.6	7.4	7.4	10.6	10.7	10.0	10.0
3	6.0	6.0	6.5	6.5	5.3	5.3	5.0	5.0	11.1	11.1	10.6	10.7	10.9	10.9	10.6	10.6
4	6.6	6.7	5.5	5.5	5.7	5.9	6.8	6.8	10.9	11.0	9.9	9.9	11.4	11.4	11.5	11.5
5	5.2	5.2	3.4	3.6	5.5	5.6	4.0	4.0	10.3	10.6	9.3	9.4	10.1	10.3	10.7	10.8
6	5.3	5.4	5.1	5.1	4.1	4.2	3.9	3.9	10.9	11.2	10.8	10.9	9.0	9.1	8.1	8.5
7	6.2	6.3	4.0	4.0	4.2	4.2	4.9	5.1	10.7	11.0	7.8	8.0	8.1	8.1	10.0	10.0
8	6.2	6.2	5.1	5.5	5.1	5.4	4.4	4.4	10.0	10.6	10.4	10.8	10.4	10.7	8.3	8.3
9	4.9	5.3	6.3	6.4	4.9	4.9	5.3	5.3	10.3	10.4	11.4	11.4	9.5	9.9	9.9	9.9
10	6.3	6.7	3.1	3.4	5.8	5.9	4.5	4.6	10.4	10.8	7.8	8.1	11.2	11.5	9.1	9.1

and $\mathrm{LL}^*_{n,M}$. Under alternatives A_2 and A_3, similar conclusions have been reached. Note that the differences in empirical powers between our two general approaches were smaller when $n = 2000$, at least in our experiments. Under alternative A_4, causality in variance in both directions occurred, with a four period lag delay. In that situation, the global test statistics displayed no power when $M \leq 3$. Under that alternative, $M = 6$ offered the best empirical powers for the test statistics.

In Tables 12-15, the power performances of the test statistics at individual lags are presented. As for the global test statistics, we observed that the test statistics $\mathrm{ER}_n(k)$ and

TABLE 7. Empirical levels, model L_3, nominal levels 5% and 10%, $n = 1000, 2000$, for the test statistics at individual lags $\mathrm{ER}_n(k)$, $\mathrm{ER}_n^*(k)$, $\mathrm{LL}_n(k)$ and $\mathrm{LL}_n^*(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
k	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
-10	6.8	7.1	5.6	5.8	6.0	6.1	5.1	5.1	10.7	10.8	10.4	10.4	11.4	11.6	11.1	11.1
-9	5.4	5.5	4.8	5.2	4.5	4.6	4.3	4.5	9.1	9.8	8.2	8.4	9.9	10.0	8.9	8.9
-8	5.6	5.7	5.9	5.9	5.6	5.6	4.8	4.8	10.7	11.1	10.3	10.3	8.8	9.0	8.7	8.8
-7	4.6	4.7	4.4	4.5	5.6	5.6	5.4	5.4	10.3	10.4	8.9	9.1	11.1	11.2	10.3	10.3
-6	6.1	6.4	4.3	4.3	4.4	4.6	5.1	5.1	10.8	10.9	9.8	10.0	9.2	9.2	9.1	9.3
-5	4.5	4.6	4.9	5.0	4.0	4.1	4.9	4.9	9.3	9.4	9.1	9.2	7.5	7.6	10.2	10.4
-4	5.9	6.0	4.4	4.4	5.7	5.8	4.2	4.3	10.7	10.7	10.2	10.3	10.8	11.1	8.9	8.9
-3	5.7	5.8	5.4	5.4	4.1	4.1	5.4	5.5	10.8	10.8	11.6	11.7	8.0	8.0	11.8	11.8
-2	5.1	5.1	5.1	5.1	5.6	5.6	5.1	5.1	9.9	9.9	9.0	9.0	9.0	9.0	11.3	11.4
-1	5.1	5.1	4.8	4.8	4.6	4.6	5.2	5.2	10.8	10.8	10.3	10.3	10.1	10.1	10.8	10.9
0	6.2	6.2	3.3	3.3	4.9	4.9	5.2	5.2	9.7	9.7	8.9	8.9	9.8	9.8	9.7	9.7
1	6.7	6.7	4.9	5.0	5.6	5.6	5.0	5.0	11.5	11.5	10.3	10.4	10.7	10.7	10.6	10.6
2	5.1	5.2	3.2	3.2	5.9	6.0	5.2	5.2	10.0	10.1	8.4	8.4	11.2	11.3	9.6	9.6
3	5.8	5.8	6.2	6.3	5.6	5.6	5.3	5.3	11.0	11.0	10.6	10.7	11.7	11.7	10.2	10.3
4	6.2	6.3	5.0	5.1	5.9	6.0	7.0	7.0	11.4	11.6	10.9	10.9	11.5	11.5	12.3	12.3
5	5.6	5.7	4.0	4.0	5.8	5.9	3.9	3.9	10.6	10.8	8.9	8.9	10.1	10.3	9.5	9.5
6	5.3	5.7	5.3	5.5	4.2	4.3	4.2	4.2	10.4	10.7	10.4	10.4	8.6	8.6	8.2	8.2
7	5.9	6.0	4.1	4.1	4.1	4.1	5.0	5.0	10.3	10.7	8.9	9.0	8.7	8.8	10.4	10.4
8	5.6	5.8	5.3	5.3	4.7	4.7	4.4	4.5	10.3	10.6	9.7	9.9	10.4	10.5	8.3	8.5
9	4.8	5.0	5.5	5.7	5.2	5.4	5.0	5.0	10.8	10.9	10.9	10.9	9.9	10.0	9.3	9.4
10	6.1	6.7	3.3	3.4	5.4	5.5	4.3	4.4	10.7	10.9	7.8	7.8	10.4	10.7	8.9	8.9

$\mathrm{ER}_n^*(k)$ were more powerful than $\mathrm{LL}_n(k)$ and $\mathrm{LL}_n^*(k)$. It should be noted that non trivial empirical powers have been observed for the same lags. Consequently, the Ling and Li test statistics seem to capture the same information on the causality patterns than the test statistics $\mathrm{ER}_n(k)$ or $\mathrm{ER}_n^*(k)$, at least in our experiments.

Under alternative A_1, all the empirical powers have been observed for lags $\pm1, \pm2, \pm3$, suggesting causality in variance in both directions, with moderate persistence. Under alternative A_2, it appears that the test statistics at individual lags are significant only for positive small order lags, detecting correctly causality in variance from $\{\mathbf{X}_{2t}\}$ to $\{\mathbf{X}_{1t}\}$. Alternative

TABLE 8. Empirical powers, model A_1, nominal levels 5% and 10%, $n = 1000, 2000$, for the global test statistics $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
M	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
1	26.9	27.0	20.4	20.4	44.7	44.8	33.3	33.3	36.6	36.7	30.0	30.1	57.6	57.7	45.7	45.7
2	51.9	52.3	39.9	39.9	86.9	86.9	71.2	71.3	64.3	64.3	50.1	50.1	91.9	91.9	80.9	80.9
3	53.6	53.8	39.2	39.3	87.3	87.3	72.5	72.6	65.0	65.3	51.2	51.4	91.9	91.9	81.9	82.0
4	50.8	51.4	37.0	37.2	85.2	85.7	70.4	70.7	63.0	63.2	48.2	48.3	90.8	91.0	78.4	78.4
5	47.2	47.5	34.1	34.3	81.7	81.8	67.5	67.7	57.8	58.2	46.0	46.3	88.5	88.7	76.2	76.2
6	44.3	44.7	31.8	31.8	77.3	77.5	64.0	64.1	54.4	55.1	42.9	43.0	84.9	85.1	74.2	74.4
7	40.9	42.0	28.7	28.7	73.5	73.7	60.5	60.6	52.5	53.9	39.2	39.9	82.2	82.6	72.2	72.5
8	36.7	38.4	27.5	27.6	70.6	70.9	58.2	58.3	49.7	51.2	38.4	38.6	79.8	80.1	69.5	69.6
9	36.0	37.3	26.2	26.4	67.3	68.0	56.7	57.0	47.7	48.9	36.5	37.0	77.5	78.0	69.1	69.2
10	33.7	35.5	24.8	25.1	64.7	65.6	54.2	54.5	47.2	48.4	34.6	35.2	76.3	76.7	67.5	67.7

TABLE 9. Empirical powers, model A_2, nominal levels 5% and 10%, $n = 1000, 2000$, for the global test statistics $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
M	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
1	18.5	18.5	13.4	13.4	26.3	26.3	20.6	20.6	26.8	27.0	21.7	21.7	36.6	36.6	31.9	31.9
2	31.1	31.4	23.8	23.9	58.6	58.6	45.8	45.9	42.0	42.2	35.0	35.1	68.8	69.0	57.7	57.8
3	31.3	31.6	24.6	24.6	56.1	56.1	47.7	47.8	41.9	42.0	33.7	33.8	68.1	68.1	59.1	59.1
4	30.4	30.8	23.5	23.6	52.6	52.9	45.9	45.9	41.0	41.6	33.4	33.6	65.2	65.9	56.5	56.5
5	25.7	26.5	21.4	21.5	48.9	49.5	41.4	41.6	35.9	36.3	31.2	31.9	60.9	61.3	54.0	54.0
6	23.1	23.9	20.1	20.3	44.1	45.0	39.3	39.4	35.0	36.0	29.5	29.6	58.5	58.9	50.8	50.8
7	21.7	22.4	18.7	18.9	40.8	41.2	36.7	36.8	32.7	33.5	27.7	27.7	55.2	55.9	47.9	48.5
8	21.6	21.9	18.5	18.7	39.4	39.7	34.6	34.8	30.4	32.0	25.9	26.3	52.3	53.0	45.8	45.9
9	18.8	19.7	17.1	17.4	35.9	36.9	33.1	33.2	29.0	30.8	25.7	25.9	47.6	48.7	44.4	44.6
10	18.7	20.0	16.1	16.2	33.9	34.8	31.2	31.6	29.1	30.5	24.4	24.8	46.9	48.2	42.9	43.3

TABLE 10. Empirical powers, model A_3, nominal levels 5% and 10%, $n = 1000, 2000$, for the global test statistics $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$, $M = 1, \dots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
M	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
1	20.1	20.2	16.0	16.1	32.2	32.3	25.0	25.1	29.5	29.7	24.6	24.6	43.9	44.0	34.4	34.4
2	37.5	37.7	27.1	27.3	67.5	67.6	53.8	53.8	49.0	49.2	39.9	40.0	78.9	79.1	66.1	66.1
3	38.8	39.3	28.8	29.1	70.1	70.4	55.8	55.9	50.5	50.8	39.8	39.9	79.6	79.7	66.1	66.2
4	36.8	37.5	26.8	26.8	65.7	66.1	53.7	53.7	48.4	48.7	37.8	38.0	77.5	77.5	64.8	64.8
5	32.1	32.3	24.0	24.5	61.3	61.7	50.1	50.1	45.4	45.9	34.7	34.8	72.6	72.8	62.0	62.0
6	30.3	31.0	22.2	22.3	56.5	56.8	45.8	46.0	41.7	42.9	32.8	33.2	68.3	68.7	58.9	59.1
7	27.5	28.2	19.1	19.6	52.5	53.0	43.2	43.5	39.6	40.4	30.7	30.7	65.1	65.9	55.8	55.9
8	26.9	27.9	19.3	19.6	47.6	48.0	39.8	40.1	37.4	38.4	29.0	29.1	62.2	62.5	52.7	53.0
9	23.7	24.9	18.7	18.9	45.5	46.1	38.0	38.3	35.6	37.2	27.8	28.3	58.3	59.4	52.2	52.3
10	24.2	25.4	18.6	19.2	42.8	43.8	37.1	37.4	34.0	36.6	27.0	27.6	56.3	57.3	50.3	50.6

TABLE 11. Empirical powers, model A_4, nominal levels 5% and 10%, $n = 1000, 2000$, for the global test statistics $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$, $M = 1, \dots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
M	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
1	5.5	5.5	4.0	4.0	4.4	4.4	5.3	5.3	9.6	9.6	8.3	8.3	8.3	8.5	10.5	10.5
2	4.8	4.8	3.6	3.7	4.6	4.6	5.9	5.9	8.9	9.0	6.9	6.9	9.4	9.5	10.7	10.7
3	5.1	5.2	4.8	4.8	4.9	4.9	3.9	4.0	9.2	9.2	9.1	9.1	8.4	8.4	9.6	9.6
4	14.9	15.3	14.1	14.2	29.2	29.2	25.3	25.3	24.4	24.7	21.8	22.2	41.6	42.1	36.9	37.0
5	33.2	34.0	27.1	27.1	68.1	68.5	54.2	54.3	46.0	47.2	37.8	38.0	77.2	77.3	66.1	66.4
6	38.6	39.5	30.2	30.4	73.7	73.7	60.6	60.8	51.3	52.4	42.3	42.5	80.9	81.4	71.1	71.3
7	37.5	38.5	29.2	29.7	71.5	72.2	59.2	59.6	50.0	51.0	41.2	41.5	80.4	80.6	71.7	71.8
8	36.9	38.5	27.0	27.5	68.5	68.9	56.6	57.1	48.9	49.9	39.6	40.2	78.6	79.1	68.8	69.1
9	33.8	35.5	26.7	26.9	65.0	65.8	52.7	52.9	46.5	48.2	36.4	37.1	75.9	76.5	67.4	67.5
10	33.0	34.8	24.6	25.2	61.6	62.4	51.9	52.2	44.8	46.1	36.2	36.5	73.8	74.4	64.3	64.8

TABLE 12. Empirical powers, model A_1, nominal levels 5% and 10%, $n = 1000, 2000$, for the test statistics at individual lags $\mathrm{ER}_n(k)$, $\mathrm{ER}_n^*(k)$, $\mathrm{LL}_n(k)$ and $\mathrm{LL}_n^*(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
k	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
-10	6.3	6.4	5.6	6.0	6.0	6.2	4.9	4.9	9.9	10.3	10.6	10.6	11.5	11.9	10.8	10.8
-9	5.3	5.3	5.1	5.1	4.9	4.9	4.8	4.8	9.4	9.4	8.7	8.8	9.9	10.4	8.9	9.1
-8	5.2	5.4	5.6	5.6	5.4	5.4	4.4	4.6	9.5	9.8	10.2	10.3	9.0	9.2	10.1	10.3
-7	5.1	5.4	4.6	4.6	6.6	6.7	5.6	5.8	10.0	10.3	8.3	8.6	11.5	11.7	11.2	11.2
-6	6.0	6.2	4.6	4.8	4.7	4.8	5.3	5.3	11.0	11.2	10.1	10.1	8.9	8.9	9.2	9.2
-5	5.4	5.4	5.0	5.0	5.6	5.6	5.4	5.4	9.4	9.6	9.5	9.5	10.2	10.5	11.4	11.4
-4	7.5	7.7	5.5	5.5	8.4	8.4	7.2	7.2	12.9	13.4	10.9	10.9	15.1	15.4	11.6	11.6
-3	13.7	13.7	12.0	12.2	20.2	20.3	17.4	17.4	22.6	23.0	19.4	19.4	29.0	29.0	25.3	25.3
-2	30.9	30.9	26.6	26.6	57.1	57.1	47.4	47.4	41.7	41.9	35.2	35.3	67.6	67.6	57.6	57.6
-1	23.2	23.3	18.0	18.1	35.7	35.7	29.1	29.1	32.5	32.6	24.5	24.5	46.9	46.9	40.0	40.0
0	4.9	4.9	3.7	3.7	5.3	5.3	5.3	5.3	10.3	10.3	8.9	8.9	9.7	9.7	9.8	9.8
1	22.9	23.0	18.3	18.3	37.0	37.1	26.7	26.7	32.0	32.0	26.8	26.8	49.1	49.1	36.5	36.5
2	30.7	30.8	25.7	25.7	55.0	55.0	45.0	45.0	41.8	42.1	35.5	35.5	64.2	64.3	57.3	57.3
3	13.0	13.1	11.0	11.0	18.1	18.5	15.9	16.0	21.0	21.1	18.2	18.2	27.2	27.3	23.9	23.9
4	8.0	8.2	6.6	6.7	10.0	10.1	9.3	9.5	14.1	14.5	12.2	12.2	16.9	16.9	15.9	15.9
5	5.7	5.7	5.0	5.0	7.0	7.1	5.1	5.1	11.1	11.6	10.1	10.2	12.1	12.2	10.6	10.6
6	5.2	5.5	5.1	5.1	5.0	5.1	3.8	4.0	11.1	11.4	11.2	11.3	10.4	10.5	9.3	9.3
7	5.9	6.1	4.1	4.1	4.5	4.5	5.0	5.0	10.7	11.1	8.2	8.4	8.7	8.7	10.1	10.1
8	6.3	6.7	5.0	5.1	5.6	5.6	4.7	4.7	10.9	11.4	9.3	9.6	10.1	10.3	8.5	8.6
9	5.3	5.7	5.8	5.9	6.2	6.3	5.4	5.4	10.3	10.8	10.5	10.8	9.8	10.1	10.7	11.0
10	5.6	6.0	3.8	3.9	5.9	6.1	4.6	4.6	10.8	11.3	8.8	8.8	11.2	11.6	9.2	9.4

A_3 is similar to alternative A_1, but the causality in variance dynamics are different between the two time series. From our simulation results, larger empirical powers are displayed for negative lags than for positive lags, suggesting stronger volatility spillover from $\{\mathbf{X}_{1t}\}$ to $\{\mathbf{X}_{2t}\}$. Given the order of magnitude of δ_{12} and γ_{12} comparatively to δ_{21} and γ_{21} for that DGP, we conclude that our test statistics are capable of describing these kind of dynamics. Finally, the causality in variance dynamic under alternative A_4 has a four period lag delay; the proposed test statistics detect causality in variance in both directions, and the larger empirical powers occurred for lags $\pm 4, \pm 5$ and ± 6. Overall, the test statistics at individual

TABLE 13. Empirical powers, model A_2, nominal levels 5% and 10%, $n = 1000, 2000$, for the test statistics at individual lags $\mathrm{ER}_n(k)$, $\mathrm{ER}_n^*(k)$, $\mathrm{LL}_n(k)$ and $\mathrm{LL}_n^*(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
k	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
-10	6.4	6.7	5.9	6.0	6.2	6.2	4.8	4.9	10.4	10.9	10.4	10.7	11.5	11.6	11.6	11.7
-9	5.2	5.7	4.6	4.9	4.8	4.9	5.0	5.0	9.6	9.8	9.2	9.2	10.5	10.9	8.1	8.1
-8	5.5	5.6	5.5	5.6	5.6	5.6	4.4	4.4	9.6	9.8	11.1	11.2	9.1	9.2	9.8	9.8
-7	4.8	4.8	4.8	4.8	5.8	6.0	5.7	5.7	10.2	10.5	9.0	9.1	11.4	11.6	11.5	11.7
-6	5.9	5.9	4.4	4.4	4.5	4.6	4.9	4.9	10.5	10.6	10.4	10.5	7.8	7.9	9.5	9.5
-5	4.0	4.2	4.5	4.5	4.5	4.6	4.8	4.8	8.9	9.4	8.8	8.9	7.8	7.8	10.2	10.2
-4	4.9	5.1	4.4	4.4	5.7	5.7	4.3	4.3	10.0	10.0	9.5	9.6	10.4	10.4	9.6	9.6
-3	6.2	6.3	5.8	5.8	4.8	4.8	5.7	5.7	10.6	10.8	10.9	10.9	8.5	8.5	11.7	11.7
-2	5.3	5.5	5.2	5.2	5.5	5.5	5.1	5.1	10.2	10.2	9.2	9.2	9.6	9.7	11.2	11.2
-1	5.1	5.1	5.0	5.0	4.9	4.9	5.2	5.2	10.6	10.6	9.3	9.3	9.7	9.8	10.6	10.6
0	5.5	5.5	3.6	3.6	5.0	5.0	5.2	5.2	10.2	10.2	8.7	8.7	9.5	9.5	9.8	9.8
1	27.2	27.2	21.1	21.2	44.3	44.3	31.9	31.9	36.0	36.2	29.4	29.4	55.4	55.4	43.6	43.6
2	35.7	36.0	30.1	30.2	62.5	62.7	52.7	52.7	48.3	48.3	40.4	40.4	70.9	70.9	63.7	63.7
3	14.5	14.8	12.2	12.3	20.2	20.4	17.1	17.1	22.3	22.6	19.9	19.9	29.2	29.3	25.8	25.9
4	9.1	9.3	6.8	6.8	9.5	9.8	9.0	9.0	15.1	15.3	13.1	13.2	16.1	16.3	16.8	16.8
5	5.5	5.5	4.7	4.7	6.9	6.9	4.9	4.9	11.5	11.6	10.0	10.1	12.1	12.2	11.5	11.6
6	5.3	5.5	4.9	5.1	5.0	5.2	4.1	4.1	10.9	11.4	11.2	11.4	8.9	9.2	8.6	8.6
7	5.9	6.1	4.5	4.5	4.4	4.4	5.5	5.7	9.6	10.1	8.4	8.6	8.6	8.6	9.8	9.8
8	5.6	6.3	4.8	4.8	4.9	5.1	4.8	4.8	11.0	11.1	9.8	9.9	9.9	9.9	8.6	8.6
9	5.2	5.4	5.7	5.7	4.9	4.9	5.0	5.1	10.4	11.1	11.1	11.4	10.1	10.4	10.1	10.2
10	6.1	6.2	3.4	3.5	5.8	5.8	4.3	4.3	10.4	10.9	8.0	8.2	10.7	10.8	8.5	9.0

lags provide important additional insight on the nature of the causality in variance. In practice, the one-lag test statistics should be interesting complements to the portmanteau test procedures. The next section illustrates the proposed methods with real data.

5. APPLICATION

To assess the ability of the proposed test statistics to detect causality in variance, we utilize financial data. The sample consists of two bivariate time series. The first time series is composed of daily observations from the North American market: the Canadian S&P/TSX

TABLE 14. Empirical powers, model A_3, nominal levels 5% and 10%, $n = 1000, 2000$, for the test statistics at individual lags $\mathrm{ER}_n(k)$, $\mathrm{ER}_n^*(k)$, $\mathrm{LL}_n(k)$ and $\mathrm{LL}_n^*(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
k	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
-10	6.3	6.6	6.0	6.1	6.4	6.4	5.1	5.1	9.8	10.6	10.4	10.8	11.3	11.4	10.8	11.0
-9	5.5	5.5	4.8	5.0	4.5	4.7	4.9	4.9	9.3	9.7	8.6	8.8	10.0	10.1	8.9	9.0
-8	5.1	5.2	5.6	5.6	5.6	5.6	4.7	4.8	10.0	10.1	10.1	10.2	8.8	9.1	10.0	10.1
-7	5.1	5.3	4.3	4.4	6.6	6.6	6.1	6.1	10.1	10.4	8.8	9.0	11.6	11.7	10.9	11.0
-6	6.3	6.7	4.3	4.4	4.6	4.6	5.4	5.5	11.3	11.4	10.2	10.2	8.8	8.9	9.1	9.3
-5	5.0	5.0	4.9	4.9	5.3	5.4	5.4	5.4	9.3	9.3	9.6	9.6	9.7	9.7	11.3	11.3
-4	7.2	7.6	5.5	5.5	7.6	7.6	6.8	6.8	12.8	13.2	10.7	10.7	14.1	14.1	11.4	11.4
-3	12.6	13.0	11.1	11.1	18.7	18.7	16.4	16.4	20.9	21.0	18.6	18.7	26.5	26.5	23.8	23.9
-2	28.6	28.9	24.3	24.3	52.4	52.5	42.9	43.0	38.6	38.6	32.5	32.8	61.8	62.0	53.4	53.4
-1	21.2	21.2	17.0	17.0	32.6	32.6	26.3	26.3	30.8	30.9	23.4	23.4	42.8	42.8	36.4	36.4
0	5.1	5.1	3.8	3.8	5.3	5.3	5.2	5.2	10.5	10.5	8.7	8.7	10.1	10.1	9.9	9.9
1	15.0	15.2	12.3	12.3	20.9	20.9	14.1	14.1	22.7	22.8	19.0	19.0	28.8	28.8	23.8	23.9
2	18.4	18.5	13.7	13.8	31.1	31.3	25.5	25.7	26.4	26.6	23.0	23.0	43.0	43.1	36.4	36.4
3	9.3	9.6	8.4	8.4	11.9	11.9	9.4	9.4	16.1	16.2	14.7	14.8	19.6	19.8	16.8	16.8
4	7.5	7.8	5.9	5.9	8.5	8.5	7.8	7.8	12.5	12.8	11.6	11.6	14.1	14.1	13.6	13.6
5	5.4	5.6	4.4	4.4	6.8	6.8	4.7	4.7	10.8	11.0	9.3	9.4	11.6	11.7	10.5	10.5
6	5.0	5.2	4.9	5.0	4.6	4.6	3.9	4.0	10.7	11.1	10.9	10.9	9.9	10.0	9.3	9.4
7	5.9	6.0	4.2	4.2	4.3	4.4	4.9	4.9	10.2	10.7	8.3	8.5	8.2	8.4	9.8	9.8
8	6.4	6.7	5.1	5.1	5.6	5.8	4.6	4.6	11.4	11.4	9.0	9.3	10.2	10.2	8.6	8.6
9	5.8	6.3	5.8	5.9	5.8	5.9	5.3	5.3	10.8	10.9	10.7	11.0	9.8	9.8	10.3	10.3
10	5.6	6.1	3.6	3.6	6.0	6.2	4.3	4.3	10.7	10.9	8.6	8.7	10.9	11.1	9.2	9.2

composite index, which represents an index of the stock (equity) prices of the largest companies on the Toronto Stock Exchange, and the well-known American S&P 500 stock index. We denote these time series $\{\text{S\&P-TSX}_t, t = 1, \ldots, n\}$ and $\{\text{S\&P500}_t, t = 1, \ldots, n\}$, respectively. The second bivariate time series includes two European indexes: the Swiss Market Index (SMI), which is composed of the largest and most liquid stocks on the Swiss stock market, and the main index of the *Place de Paris*, the French CAC 40. The data are observed on a daily basis (from Monday to Friday), beginning on January 3, 2007, and ending on December 31, 2009. The resulting time series are noted $\{\mathrm{SMI}_t, t = 1, \ldots, n\}$ and

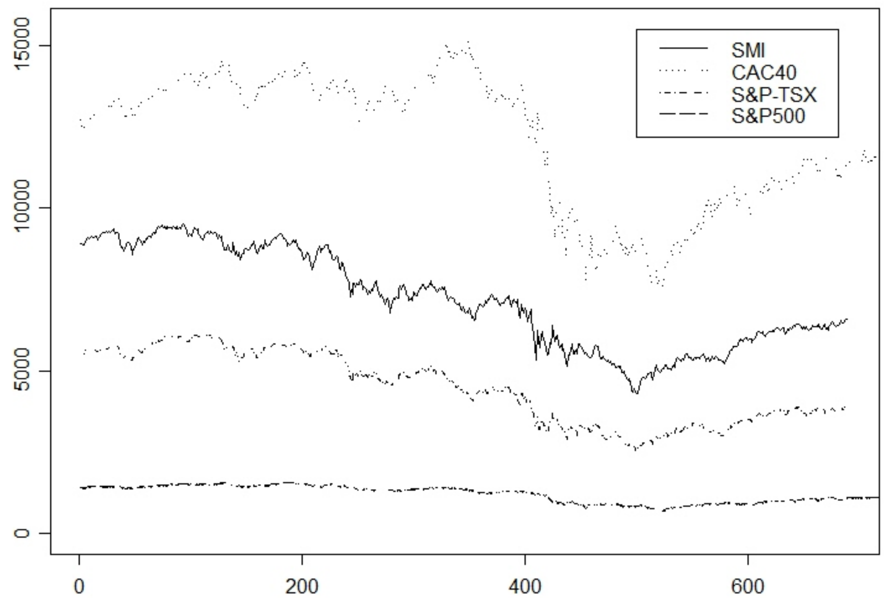
TABLE 15. Empirical powers, model A_4, nominal levels 5% and 10%, $n = 1000, 2000$, for the test statistics at individual lags $\mathrm{ER}_n(k)$, $\mathrm{ER}^*_n(k)$, $\mathrm{LL}_n(k)$ and $\mathrm{LL}^*_n(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
k	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
-10	6.2	6.2	4.9	5.2	6.9	7.1	5.4	5.4	10.8	11.1	10.9	11.2	11.6	11.8	10.8	10.9
-9	4.9	5.2	4.6	4.6	5.0	5.1	4.4	4.4	9.8	10.0	8.9	9.2	11.0	11.2	10.1	10.2
-8	6.5	6.5	6.6	6.6	6.5	6.7	5.1	5.2	10.9	11.1	10.9	10.9	11.3	11.5	9.6	9.6
-7	8.0	8.4	5.8	5.9	9.9	10.1	8.3	8.3	13.7	14.3	10.8	10.8	15.8	16.1	13.4	13.5
-6	14.4	14.6	12.6	12.7	19.5	19.5	18.6	18.6	22.1	22.6	18.4	18.4	29.7	30.2	26.4	26.4
-5	34.1	34.5	30.2	30.5	62.7	62.8	50.7	50.9	44.7	44.8	41.1	41.2	71.3	71.3	61.2	61.3
-4	19.5	20.0	14.9	14.9	38.5	38.7	28.8	29.0	29.1	29.7	23.1	23.1	51.7	52.1	39.1	39.1
-3	6.0	6.0	5.3	5.3	4.5	4.5	6.7	6.8	10.4	10.4	10.9	10.9	8.4	8.5	12.5	12.5
-2	5.8	5.8	5.3	5.3	5.2	5.2	5.4	5.4	9.5	9.6	9.2	9.2	8.9	8.9	10.7	10.7
-1	5.6	5.6	5.1	5.1	4.3	4.3	5.9	5.9	11.0	11.3	10.0	10.1	8.8	8.9	11.4	11.4
0	4.9	4.9	3.5	3.5	4.5	4.5	5.0	5.0	8.4	8.4	8.5	8.5	10.2	10.2	9.7	9.7
1	6.5	6.5	5.3	5.3	4.6	4.6	4.8	4.8	11.5	11.5	10.4	10.4	10.7	10.7	9.9	9.9
2	5.3	5.3	3.4	3.4	5.0	5.0	5.7	5.7	8.9	9.0	8.0	8.1	10.8	10.8	10.1	10.2
3	5.8	6.0	5.5	5.5	4.8	4.8	5.0	5.0	11.8	12.1	11.5	11.5	10.7	10.7	9.3	9.3
4	28.2	28.4	20.2	20.2	44.6	44.9	32.5	32.7	37.9	38.2	30.4	30.5	54.6	54.6	43.1	43.2
5	24.9	25.5	18.0	18.1	44.6	44.9	31.3	31.3	34.6	35.1	27.5	27.8	57.1	57.1	41.9	42.0
6	11.1	11.5	9.5	9.6	15.1	15.3	12.7	12.7	18.2	18.6	16.5	16.6	24.5	24.5	20.3	20.3
7	7.7	7.9	4.7	4.7	8.7	8.8	7.9	8.2	13.3	13.7	9.2	9.3	14.1	14.4	14.7	14.7
8	6.2	6.7	5.1	5.1	5.6	5.9	5.1	5.1	11.9	12.3	11.1	11.2	11.6	11.7	10.2	10.2
9	5.8	6.2	5.4	5.4	5.6	5.8	5.6	5.7	10.8	11.1	10.8	10.9	10.7	10.9	9.7	9.7
10	5.9	6.2	3.0	3.1	6.0	6.4	4.8	4.8	10.8	10.9	8.4	8.5	11.7	12.2	9.0	9.1

$\{\mathrm{CAC40}_t, t = 1, \ldots, n\}$, respectively. From the 784 days, missing data due to holidays were removed before the analysis. The final data set included $n = 724$ observations. Figure 1 displays the original data.

A descriptive analysis of each individual time series suggested non-stationarity. Daily returns have been calculated, using a logarithm transformation of the original variables, and applying the difference filter $1 - B$ on the data. The daily returns now appeared stationary. The transformed time series are presented in Figure 2. Thus, the time series used in the

Figure 1. Swiss Market Index, CAC 40 Index, S &P - TSX Composite Index and S & P 500 Index from January 2007 to December 2009.



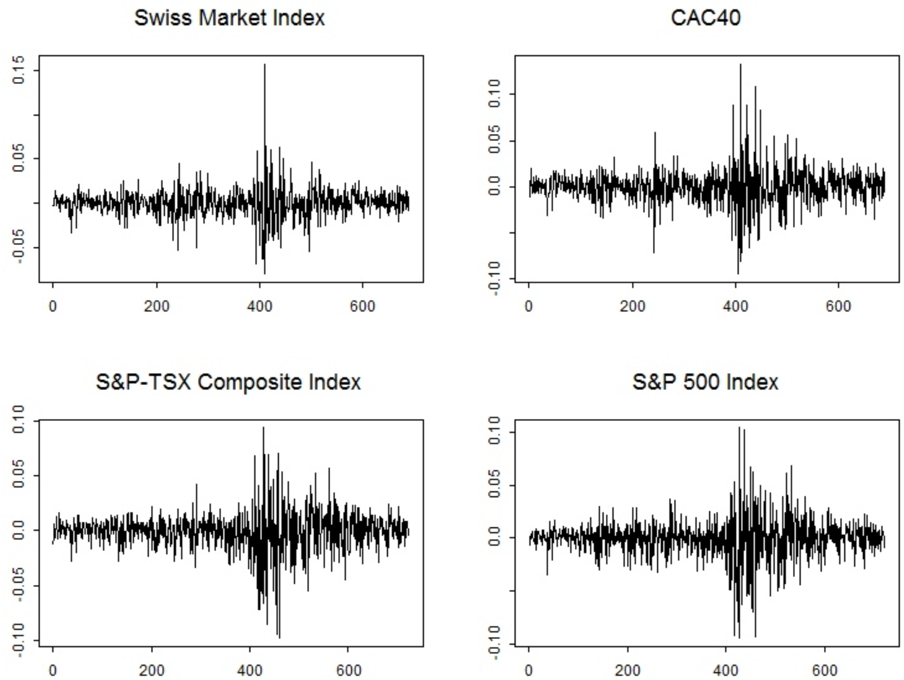
analysis are defined as follows:

$$\mathbf{X}_{1t} = \begin{pmatrix} X_{1t}(1) \\ X_{1t}(2) \end{pmatrix} = \begin{pmatrix} (1-B)\log(\text{S\&P-TSX}_t) \\ (1-B)\log(\text{S\&P500}_t) \end{pmatrix},$$

$$\mathbf{X}_{2t} = \begin{pmatrix} X_{2t}(1) \\ X_{2t}(2) \end{pmatrix} = \begin{pmatrix} (1-B)\log(\text{SMI}_t) \\ (1-B)\log(\text{CAC40}_t) \end{pmatrix}.$$

In order to test for causality in variance between the indexes from North America with those coming from Europe, each bivariate time series is fitted separately. We recall that our approach does not involve a global model for the two markets; it appears simple and natural to model each market separately, and to investigate causality in mean and causality in variance with appropriate test statistics. As in Section 5, the S-PLUS software with the `Finmetrics` module is used to adjust the time series. We postulated models in the VARMA-GARCH family. For the conditional mean, we examined several competing formulations: white noise models, vector autoregressive VAR(p) models with small values of p, vector moving-average VMA(q) models with small values of q, and finally VARMA models. A DVEC(1,1) model (see, e.g., Zivot and Wang (2003)) for the conditional variance has been

FIGURE 2. Swiss Market Index, CAC 40 Index, S &P - TSX Composite Index and S & P 500 Index after a logarithm transformation and a first-order differentiation.



considered. We retained a VMA(1) model for returns from the North American market, and a white noise model for the European returns. More precisely, we fitted the following models:

$$\mathbf{X}_{1t} = (\mathbf{I}_2 + \boldsymbol{\Theta} B)\boldsymbol{\epsilon}_{1t},$$
$$\mathbf{X}_{2t} = \boldsymbol{\epsilon}_{2t},$$

and $\boldsymbol{\Theta} = (\Theta_{ij})_{i,j\in\{1,2\}}$. For each model, the conditional variances are postulated to be of the following form:

$$\mathbf{V}_t = \mathbf{C} + \mathbf{A} \odot (\mathbf{X}_{1,t-1}\mathbf{X}_{1,t-1}^{\top}) + \mathbf{B} \odot \mathbf{V}_{t-1},$$

where $'\odot'$ corresponds to the Hadamard product, defined as the element-by-element multiplication. Since $\mathbf{V}_t$ represents a conditional covariance matrix, the model parameters $\mathbf{C} = (C_{ij})_{i,j\in\{1,2\}}$, $\mathbf{A} = (A_{ij})_{i,j\in\{1,2\}}$ and $\mathbf{B} = (B_{ij})_{i,j\in\{1,2\}}$ are supposed to be symmetric. The estimators are given in Table 16. Several estimators are significant at the 5% nominal level, and all the t-values are larger than one in absolute value. For each market, all the estimators of the parameters in the conditional variance functions are highly significant.

TABLE 16. Estimators of the models for the North American and the European indexes with their corresponding t-values.

	North American indexes			European indexes		
	Estimators	t-values	P-values	Estimators	t-values	P-values
$\hat{\Theta}_{11}$	-0.1302	-2.714	0.0068			
$\hat{\Theta}_{21}$	-0.0963	-2.127	0.0338			
$\hat{\Theta}_{12}$	0.1313	3.022	0.0026			
$\hat{\Theta}_{22}$	-0.0579	-1.074	0.2831			
$\hat{C}_{11}$	3.717×10^{-6}	2.958	0.0032	6.099×10^{-6}	4.273	0^+
$\hat{C}_{21}$	3.209×10^{-6}	3.171	0.0015	6.012×10^{-6}	4.042	0^+
$\hat{C}_{22}$	4.531×10^{-6}	3.422	0.0006	6.791×10^{-6}	3.967	0^+
$\hat{A}_{11}$	0.084	5.830	0^+	0.111	7.994	0^+
$\hat{A}_{21}$	0.076	6.006	0^+	0.091	7.512	0^+
$\hat{A}_{22}$	0.088	5.933	0^+	0.089	6.852	0^+
$\hat{B}_{11}$	0.897	54.238	0^+	0.855	46.171	0^+
$\hat{B}_{21}$	0.908	59.959	0^+	0.873	49.580	0^+
$\hat{B}_{22}$	0.892	49.548	0^+	0.884	53.671	0^+

Consider the test statistics at individual lags defined as follows:

$$\mathrm{H}_n^{(ij)}(k) = n\mathrm{tr}\{\mathbf{C}_{\hat{\boldsymbol{\eta}}_i\hat{\boldsymbol{\eta}}_j}(k)\mathbf{C}^{-1}_{\hat{\boldsymbol{\eta}}_j\hat{\boldsymbol{\eta}}_j}(0)\mathbf{C}^{\top}_{\hat{\boldsymbol{\eta}}_i\hat{\boldsymbol{\eta}}_j}(k)\mathbf{C}^{-1}_{\hat{\boldsymbol{\eta}}_i\hat{\boldsymbol{\eta}}_i}(0)\}, \tag{22}$$

$$\mathrm{HM}_n^{(ij)}(k) = n\left(\frac{n}{n-k}\right)\mathrm{tr}\{\mathbf{C}_{\hat{\boldsymbol{\eta}}_i\hat{\boldsymbol{\eta}}_j}(k)\mathbf{C}^{-1}_{\hat{\boldsymbol{\eta}}_j\hat{\boldsymbol{\eta}}_j}(0)\mathbf{C}^{\top}_{\hat{\boldsymbol{\eta}}_i\hat{\boldsymbol{\eta}}_j}(k)\mathbf{C}^{-1}_{\hat{\boldsymbol{\eta}}_i\hat{\boldsymbol{\eta}}_i}(0)\}. \tag{23}$$

When $i = j \in \{1,2\}$, the test statistics $\mathrm{H}_n^{(ii)}(k)$ and $\mathrm{HM}_n^{(ii)}(k)$, $i = 1,2$, are useful in order to quantify the remaining dependence in the standardized residuals at each lag k (see, e.g., Hosking (1980) and Ansley and Newbold (1979)). In Table 17, these test statistics have been calculated on the standardized residuals of each model for $k = 1,\ldots,12$. Since $d_1 = d_2 = 2$, the approximate critical values of $\mathrm{H}_n^{(ii)}(k)$ and $\mathrm{HM}_n^{(ii)}(k)$ are taken from the χ^2_4 distribution; at the 5% and 1% nominal levels, the quantiles are 9.5 and 13.3, respectively. Except maybe at lag one for the European market, the results suggest that the models are reasonably well fitted. Note that the Hosking portmanteau test statistics can be deduced from these test statistics and the same general conclusions are obtained. A residual analysis also suggested that the chosen multivariate models were reasonably well fitted. In checking the adequacy of univariate time series models, Hong (2001) used Box-Pierce-Ljung test statistics on standardized residuals in order to validate univariate ARMA-GARCH models.

TABLE 17. Test statistics at individual lags applied on the standardized residuals.

	$\mathrm{H}_n^{(ii)}(k)$		$\mathrm{HM}_n^{(ii)}(k)$	
k	$\mathbf{X}_{1t}$	$\mathbf{X}_{2t}$	$\mathbf{X}_{1t}$	$\mathbf{X}_{2t}$
1	0.7549	12.0028	0.7731	12.0527
2	4.6150	7.7570	4.7243	8.0935
3	4.7946	4.4588	4.9030	4.5681
4	3.5571	1.2276	3.6461	1.2379
5	0.8209	2.0551	0.8334	2.0814
6	1.5640	1.5452	1.6119	1.5827
7	1.2945	3.2998	1.3342	3.4570
8	0.3778	1.9976	0.3903	2.0455
9	4.8130	4.9701	4.9837	5.1216
10	4.8083	1.6018	4.9059	1.6347
11	10.6681	5.8460	11.0520	6.0312
12	3.0415	3.5822	3.1532	3.7328

Before to test for causality in variance, causality in mean relations have been investigated. We calculated the test statistics at individual lags of El Himdi and Roy (1997) applied on the standardized residuals. These test statistics are defined as (22) and (23) and correspond to $\mathrm{H}_n^{(ij)}(k)$ and $\mathrm{HM}_n^{(ij)}(k)$, with $i \neq j$. That practice is similar to the univariate analysis described in Cheung and Ng (1996, Section 4), see also Hong (2001, pp. 206-207). The results are presented in Table 19; they suggest significant causality in mean relations at lags 0 and -1. Note that causality in mean is also found at the high order lag 9 at the 5% nominal level (but not at the 1% nominal level); due to the large number of test statistics, it is not excluded that a test statistic rejects the null when in fact it is true. We generalize the approach of Cheung and Ng (1993) in our context, and we also adjust augmented models. We fitted the following augmented models:

$$\begin{aligned} \mathbf{X}_{1t} &= (\mathbf{I}_2 + \mathbf{\Theta}^{(1)}B)\boldsymbol{\epsilon}_{1t}, \\ (\mathbf{I}_2 - \mathbf{\Phi}^{(2)}B)\mathbf{X}_{2t} &= \mathbf{\Lambda}_0^{(2)}\mathbf{X}_{1t} + \mathbf{\Lambda}_1^{(2)}\mathbf{X}_{1,t-1} + (\mathbf{I}_2 + \mathbf{\Theta}^{(2)}B)\boldsymbol{\epsilon}_{2t}, \end{aligned}$$

where we retained the same formulations for the conditional variances. Based on the augmented models, the causality in mean relations have been investigated between the two time series. The results are displayed in Table 19. The inclusion of additional lags and exogenous variables reduced considerably the causality in mean relations, except maybe at lags -5 and

9. Note that we investigated smaller models for $\{\mathbf{X}_{2t}\}$ but these models did not removed completely the causality in mean dynamics between the time series; larger models seemed necessary. The estimators of the augmented model for $\{\mathbf{X}_{2t}\}$ are given in Table 18.

Having removed considerably causality in mean relations, we now apply the new test statistics for checking causality in variance. We report the results for the test statistics at individual lags $\mathrm{ER}_n^*(k)$ and $\mathrm{LL}_n^*(k)$, $k = -10, \ldots, 10$, and also for the global test statistics $\mathrm{ER}_{n,M}^*$ and $\mathrm{LL}_{n,M}^*$, $k = 1, \ldots, 10$. From Table 20, all the P-values of the global test statistics are highly significant, at any reasonable significance level. It is useful to investigate causality in variance relations, and the test statistics at individual lags serve well that purpose. From Table 21, the conclusions between the two class of test statistics seem generally in accordance: Causality in variance is found for negatives lags, suggesting that the North American market causes the European market. Note that the P-value of $\mathrm{ER}_n^*(-2)$ is not significant at the 5% level, but the one of $\mathrm{LL}_n^*(-2)$ is slightly significant at that nominal level (but it is not significant at the 1% level). The P-value associated to $\mathrm{ER}_n^*(-3)$ is highly significant at any reasonable significance level, while it is slightly larger than 5% for the test statistic $\mathrm{LL}_n^*(-3)$;

Table 18. Estimators of the parameters of the augmented model for European indexes with their corresponding t-values.

	Estimators	t-values	P-values		Estimators	t-values	P-values
$\hat{\Phi}_{11}^{(2)}$	0.8374	4.87199	0^+	$\hat{C}_{1,11}^{(2)}$	3.316×10^{-6}	3.48746	0.0005
$\hat{\Phi}_{21}^{(2)}$	0.1099	0.50853	0.6112	$\hat{C}_{1,21}^{(2)}$	3.225×10^{-6}	3.70952	0.0002
$\hat{\Phi}_{12}^{(2)}$	-0.5988	-3.71688	0.0002	$\hat{C}_{1,22}^{(2)}$	4.139×10^{-6}	3.99720	0^+
$\hat{\Phi}_{22}^{(2)}$	-0.007787	-0.03926	0.9687	$\hat{A}_{1,11}^{(2)}$	0.1140	7.08866	0^+
$\hat{\Theta}_{11}^{(2)}$	-0.9885	-6.40815	0^+	$\hat{A}_{1,21}^{(2)}$	0.09915	7.49758	0^+
$\hat{\Theta}_{21}^{(2)}$	-0.1753	-0.86578	0.3869	$\hat{A}_{1,22}^{(2)}$	0.1083	7.85132	0^+
$\hat{\Theta}_{12}^{(2)}$	0.3657	2.56920	0.0104	$\hat{B}_{1,11}^{(2)}$	0.8555	40.69159	0^+
$\hat{\Theta}_{22}^{(2)}$	-0.3772	-2.03904	0.0418	$\hat{B}_{1,21}^{(2)}$	0.8585	44.77057	0^+
$\hat{\Lambda}_{0,11}^{(2)}$	0.06618	2.29225	0.0221	$\hat{B}_{1,22}^{(2)}$	0.8575	52.97661	0^+
$\hat{\Lambda}_{0,21}^{(2)}$	0.1716	4.75154	0^+				
$\hat{\Lambda}_{0,12}^{(2)}$	0.4427	16.07953	0^+				
$\hat{\Lambda}_{0,22}^{(2)}$	0.5449	16.07594	0^+				
$\hat{\Lambda}_{1,11}^{(2)}$	-0.04050	-0.91966	0.3581				
$\hat{\Lambda}_{1,21}^{(2)}$	-0.04125	-0.92451	0.3555				
$\hat{\Lambda}_{1,12}^{(2)}$	0.2809	6.01741	0^+				
$\hat{\Lambda}_{1,22}^{(2)}$	0.2822	5.70388	0^+				

Table 19. Test statistics of El Himdi and Roy (1997) for testing causality in mean. Approximate P-values are given.

	Original models		Augmented models	
k	$\mathrm{HM}_n^{(12)}(k)$	P-values	$\mathrm{HM}_n^{(12)}(k)$	P-values
-10	3.2333	0.5196	1.9558	0.7439
-9	3.3988	0.4934	3.7403	0.4423
-8	2.6224	0.6229	2.6881	0.6113
-7	2.1249	0.7128	1.8000	0.7725
-6	0.8094	0.9372	4.3490	0.3608
-5	5.0570	0.2815	9.0745	0.0593
-4	1.1318	0.8892	3.9421	0.4139
-3	2.4805	0.6481	7.6306	0.1061
-2	4.4317	0.3507	1.8949	0.7551
-1	59.4439	0^+	7.5140	0.1111
0	314.0003	0^+	0.5818	0.9651
1	0.7628	0.9434	2.9395	0.5680
2	2.1155	0.7145	1.2431	0.8710
3	1.8228	0.7683	2.6045	0.6260
4	1.4934	0.8278	2.1376	0.7105
5	2.5771	0.6309	2.8648	0.5807
6	0.9574	0.9162	1.0472	0.9026
7	1.6793	0.7945	1.3071	0.8602
8	1.8549	0.7624	1.2001	0.8781
9	12.9997	0.0113	10.2599	0.0363
10	3.9911	0.4072	3.7746	0.4374

that may be partly explained by a possible power loss of the test statistic calculated using the approach of Ling and Li (1997) for that lag. However, the general conclusions on the causality in variance dynamics are the same under both approaches.

6. DISCUSSION AND CONCLUSION

In this article, the procedures to test causality in variance developed by Cheung and Ng (1996) for univariate time series have been generalized in several directions. Relying on the

TABLE 20. Global test statistics for causality in variance $\mathrm{ER}^*_{n,M}$ and $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$. The degrees of freedom are given under the column untitled $d.f.$. The P-values are also provided.

	Test statistic $\mathrm{ER}^*_{n,M}$			Test statistic $\mathrm{LL}^*_{n,M}$		
M	$\mathrm{ER}^*_{n,M}$	$d.f.$	P-values	$\mathrm{LL}^*_{n,M}$	$d.f.$	P-values
1	48.1950	27	0.0073	18.9753	3	0.0003
2	66.3151	45	0.0210	25.9379	5	0.0001
3	96.5014	63	0.0042	29.4479	7	0.0001
4	162.7471	81	0^+	53.7310	9	0^+
5	199.4816	99	0^+	62.0635	11	0^+
6	217.3490	117	0^+	63.8579	13	0^+
7	229.4658	135	0^+	66.0332	15	0^+
8	243.7263	153	0^+	69.2014	17	0^+
9	256.6041	171	0^+	73.8730	19	0^+
10	293.8810	189	0^+	80.4530	21	0^+

work of El Himdi and Roy (1997) and Duchesne (2004), a first approach developed test statistics based on residual cross-covariance matrices of squared (standardized) residuals and cross products of (standardized) residuals. In a second approach, transformed residuals were defined as in Ling and Li (1997) for each residual vector time series, and test statistics were constructed based on the cross-correlations of these transformed residuals. New test statistics with convenient asymptotic chi-square distributions under the null hypothesis of absence of causality in variance have been proposed. In both approaches, test statistics at individual lags were developed, and also portmanteau type test statistics. The proposed methodology has been used to determine the directions of causality in variance, and appropriate test statistics were presented. Simulation results showed that the new test statistics offered satisfactory empirical properties, since the empirical levels were reasonably close to the nominal empirical levels. It seemed that the correction factor ω_k was not very important, which may be explained by the relatively large sample sizes considered in our study. The global test statistics $\mathrm{ER}_{n,M}$ and $\mathrm{ER}^*_{n,M}$ were more powerful than the test statistics $\mathrm{LL}_{n,M}$ and $\mathrm{LL}^*_{n,M}$. We observed similar conclusions for the test statistics at individual lags. However, for large n, the power differences were relatively small, at least in our empirical studies. Thus the test statistics $\mathrm{LL}_{n,M}$, $\mathrm{LL}^*_{n,M}$, $\mathrm{LL}_n(k)$ and $\mathrm{LL}^*_n(k)$ may be recommended for use in practical applications, given their inherent computational simplicity, particularly when the sample sizes are

TABLE 21. Test statistics at individual lags for detecting causality in variance $\mathrm{ER}_n^*(k)$ and $\mathrm{LL}_n^*(k)$, $|k| \leq 10$.

	Test statistic $\mathrm{ER}_n^*(k)$		Test statistic $\mathrm{LL}_n^*(k)$	
k	$\mathrm{ER}_n^*(k)$	P-values	$\mathrm{LL}n^*(k)$	P-values
-10	29.6308	0.0005	5.4231	0.0199
-9	4.9218	0.8411	0.0573	0.8109
-8	7.4099	0.5945	0.4335	0.5103
-7	6.2407	0.7156	2.1706	0.1407
-6	6.5046	0.6885	0.7406	0.3895
-5	25.0034	0.0030	7.3205	0.0068
-4	56.9884	0^+	20.9524	0^+
-3	27.3458	0.0012	3.2855	0.0699
-2	12.7446	0.1745	6.2150	0.0127
-1	13.4601	0.1429	1.0343	0.3091
0	27.3659	0.0012	16.5646	0^+
1	7.3689	0.5988	1.3764	0.2407
2	5.3755	0.8004	0.7476	0.3872
3	2.8406	0.9703	0.2245	0.6357
4	9.2572	0.4139	3.3306	0.0680
5	11.7311	0.2289	1.0120	0.3144
6	11.3628	0.2517	1.0539	0.3046
7	5.8762	0.7522	0.0047	0.9455
8	6.8506	0.6527	2.7347	0.0982
9	7.9560	0.5386	4.6143	0.0317
10	7.6461	0.5702	1.1570	0.2821

large. However, if powerful test statistics are needed, it may be better to compute $\mathrm{ER}_{n,M}$, $\mathrm{ER}_{n,M}^*$, $\mathrm{ER}_n(k)$ and $\mathrm{ER}_n^*(k)$. They may be more complicated to implement, but they are expected to be more powerful than those relying on the dependence measure of Ling and Li (1997). An application with financial data illustrated the methods. We tested for causality in mean relations in a first step, and working with augmented models, test statistics for causality in variance have been calculated. In our application, the test statistics $\mathrm{ER}_{n,M}^*$ and $\mathrm{ER}_n^*(k)$, and those based on Ling and Li's approach, gave similar general conclusions.

REFERENCES

Ansley, C. F. and Newbold, P. (1979), 'On the finite sample distribution of residual autocorrelations in autoregressive-moving average models', *Biometrika*, **66**, 547–553.

Berndt, E., Hall, B., Hall, R. and Hausman, J. (1974), 'Estimation and inference in nonlinear structural models', *Annals of Social Measurement*, Vol. 3, 653–665.

Brockwell, P. J. and Davis, R. A. (1991), *Time Series: Theory and Methods*, Second Edition, Springer-Verlag: New York.

Caporin, M. (2007), 'Variance (non) causality in multivariate GARCH', *Econometric Reviews* **26**, 1–24.

Cheug, Y.-W. and Ng, L. K. (1990), 'The dynamics of S&P 500 index and S&P 500 futures intraday price volatilities', *Review of Futures Markets* **9**, 458–486.

Cheug, Y.-W. and Ng, L. K. (1996), 'A causality-in-variance test and its application to financial market prices', *Journal of Econometrics* **72**, 33–48.

Comte, F. and Lieberman, O. (2000), 'Second-order noncausality in multivariate GARCH processes', *Journal of Time Series Analysis* **21**, 535–557.

Comte, F. and Lieberman, O. (2003), 'Asymptotic theory for multivariate GARCH processes', *Journal of Multivariate Analysis* **84**, 61–84.

Duchesne, P. (2004), 'On matricial measures of dependence in vector ARCH models with applications to diagnostic checking', *Statistics & Probability Letters* **68**, 149–160.

Duchesne, P. and Lalancette, S. (2002), 'A test for multivariate autoregressive conditional heteroskedasticity effects in vector time series models', *Les Cahiers du GERAD* G-2002-45, GERAD, August 2002.

Duchesne, P. and Lalancette, S. (2003), 'On testing for multivariate ARCH effects in vector time series models', *The Canadian Journal of Statistics* **31**, 275–292.

El Himdi, K. and Roy, R. (1997), 'Tests for noncorrelation of two multivariate ARMA time series', *The Canadian Journal of Statistics* **25**, 233–256.

Francq, C. and Zakoian, J.-M. (2010), 'QML estimation of a class of multivariate GARCH models without moment conditions on the observed process', Working Paper, MPRA Paper No. 20779, February 2010.

Granger, C. W. J., Robins, R. P. and Engle, R. F. (1986), 'Wholesale and retail prices: Bivariate time-series modeling with forecastable error variances', in Belsley, D. A. and Kuh, E., Editors, *Model Reliability*, MIT Press, Cambridge, MA, 1–17.

Hafner, C. M. and Herwartz, H. (2006), 'A Lagrange multiplier test for causality in variance', *Economics Letters* **93**, 137–141.

Hafner, C. M. and Herwartz, H. (2008), 'Testing for causality in variance using multivariate GARCH models', *Annales d'Économie et de Statistique* **89**, 215–241.

Hafner, C. M. and Preminger, A. (2009), 'On asymptotic theory for multivariate GARCH models', *Journal of Multivariate Analysis* **100**, 2044–2054.

Harville, D. A. (1997), *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag: New York.

Haugh, L. D. (1976), 'Checking the independence of two covariance-stationary time series: A univariate residual cross-correlation approach', *Journal of the American Statistical Association* **71**, 378–385.

Hong, Y. (2001), 'A test for volatility spillover with application to exchange rates', *Journal of Econometrics* **103**, 183–224.

Hosking, J. R. M. (1980), 'The multivariate portmanteau statistic', *Journal of the American Statistical Association* **75**, 602–608.

Li, W. K. (2004), *Diagnostic Checks in Time Series*, Chapman & Hall/CRC: New York.

Li, W. K. and Mak, T. K. (1994), 'On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity', *Journal of Time Series Analysis* **15**, 627–636.

Ling, S. and Li, W. K. (1997), 'Diagnostic checking of nonlinear multivariate time series with multivariate ARCH errors', *Journal of Time Series Analysis* **18**, 447–464.

Ling, S. and McAleer, M. (2003), 'Asymptotic theory for a vector ARMA-GARCH model', *Econometric Theory* **19**, 280–310.

Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag: New York.

McLeod, A. I. and Li, W. K. (1983), 'Diagnostic checking ARMA time series models using squared-residual autocorrelations', *Journal of Time Series Analysis* **4**, 269–273.

Ross, S. A. (1989), 'Information and volatility: The no-arbitrage martingale approach to timing and resolution irrelevancy', *Journal of Finance* **44**, 1–17.

Tse, Y. K. (2002), 'Residual-based diagnostics for conditional heteroscedasticity models', *Econometrics Journal* **5**, 358–373.

Tse, Y. K. and Tsui, A. K. C. (1999), 'A note on diagnosing multivariate conditional heteroscedasticity models', *Journal of Time Series Analysis* **20**, 679–691.

Wong, H. and Li, W. K. (1996), 'Distribution of the cross-correlations of squared residuals in ARIMA models', *The Canadian Journal of Statistics* **24**, 489–502.

Wong, H. and Li, W. K. (2002), 'Detecting and diagnostic checking multivariate conditional heteroscedastic time series models', *Annals of the Institute of Statistical Mathematics* **54**, 45–59.

Zivot, E. and Wang, J. (2003), *Modeling Financial Time Series with S-PLUS*, Springer-Verlag: New York.

Chapitre 3

Démonstration des principaux résultats

Dans cette partie nous présentons la preuve des deux théorèmes principaux de l'article.

3.1 Preuve de la proposition 1

On démontrera cette proposition en utilisant un développement de Taylor. En effet en effectuant un développement de Taylor de $\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2}$ nous avons :

$$\mathbf{c}_{\hat{\mathbf{Z}}_1\hat{\mathbf{Z}}_2} = \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2} + \frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}}{\partial \boldsymbol{\theta}^\top}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + \mathbf{o}_P(n^{-1/2}).$$

Dans ce qui suit, nous allons nous concentrer sur le modèle suivant :

$$\begin{aligned} \mathrm{vech}(\mathbf{V}_{it}) &= \boldsymbol{\omega}_i + \sum_{j=1}^{\infty} \mathbf{A}_{ij}\left(\boldsymbol{\Lambda}_i\right)\left\{\mathrm{vech}\big(\boldsymbol{\epsilon}_{i,t-j}\boldsymbol{\epsilon}_{i,t-j}^\top\big) - \boldsymbol{\omega}_i\right\}, && \text{i=1,2} \\ \boldsymbol{\mu}_{it} &= \sum_{j=1}^{\infty} \boldsymbol{\Phi}_{ij}\left(\boldsymbol{\beta}_j\right)\mathbf{X}_{i,t-j}. && \text{i=1,2} \end{aligned}$$

Notons que l'article énonce des résultats dans un cadre plus général mais le modèle ci-dessus permet de se convaincre que le résultat est valide pour une grande classe de modèles utiles en pratique. Sans perte de généralité on va se restreindre au cas $k \geq 0$ (le cas $k \leq 0$ se déduit facilement en faisant jouer à $\mathbf{u}_{1t}$ et $\mathbf{u}_{2t}$ les rôles symétriques) et on commencera par montrer que :

$$\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\alpha}^\top} = \mathbf{O}_P(n^{-1/2}),$$

où $\boldsymbol{\alpha}$ est l'un des coeffficients $\boldsymbol{\beta}_i\,, \boldsymbol{\omega}_i\,, \mathrm{vec}(\boldsymbol{\Lambda}_i),\ i=1,2$. Rappelons aussi que :

$$\begin{aligned}\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\alpha}^\top} &= n^{-1}\sum_{t=k+1}^{n}\frac{\partial \mathrm{vec}(\mathbf{u}_{1t}\mathbf{u}_{2,t-k}^\top)}{\partial \boldsymbol{\alpha}^\top},\\ &= n^{-1}\sum_{t=k+1}^{n}\left\{(\mathbf{I}_{d_2'}\otimes \mathbf{u}_{1t})\frac{\partial \mathbf{u}_{2,t-k}}{\partial \boldsymbol{\alpha}^\top}+(\mathbf{u}_{2,t-k}\otimes \mathbf{I}_{d_1'})\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\alpha}^\top}\right\}.\end{aligned}$$

<u>**Premier cas :**</u> $\boldsymbol{\alpha}=\boldsymbol{\beta}_1$

En utilisant la relation précédente nous trouvons :

$$\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\beta}_1^\top}=n^{-1}\sum_{t=k+1}^{n}\left(\mathbf{u}_{2,t-k}\otimes \mathbf{I}_{d_1'}\right)\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\beta}_1^\top}.$$

Or $\mathbf{u}_{1t}=\mathrm{vech}\Big(\mathbf{L}_{1t}^{-1}(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})^\top\left(\mathbf{L}_{1t}^{-1}\right)^\top\Big)-\mathrm{vech}\big(\mathbf{I}_{d_1'}\big)$. Ainsi nous obtenons :

$$\begin{aligned}\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\beta}_1^\top} &= \frac{\partial\left\{\mathrm{vech}\Big(\mathbf{L}_{1t}^{-1}(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})^\top\left(\mathbf{L}_{1t}^{-1}\right)^\top\Big)-\mathrm{vech}\big(\mathbf{I}_{d_1'}\big)\right\}}{\partial \boldsymbol{\beta}_1^\top},\\ &= \frac{\partial\left\{\mathrm{vech}\Big(\mathbf{L}_{1t}^{-1}(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})^\top\left(\mathbf{L}_{1t}^{-1}\right)^\top\Big)\right\}}{\partial \boldsymbol{\beta}_1^\top}.\end{aligned}$$

En utilisant la relation $\mathrm{vech}(\mathbf{A})=\mathbf{H}_n\mathrm{vec}(\mathbf{A})$, où $\mathbf{A}$ est une matrice symétrique de dimension $n\times n$ (voir Harville (1997, p. 354)), nous obtenons alors :

$$\begin{aligned}\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\beta}_1^\top} &= \mathbf{H}_{1d_1}\frac{\partial \mathrm{vec}\Big(\mathbf{L}_{1t}^{-1}(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})^\top\left(\mathbf{L}_{1t}^{-1}\right)^\top\Big)}{\partial \boldsymbol{\beta}_1^\top},\\ &= \mathbf{H}_{1d_1}\Bigg\{\left(\mathbf{I}_{d_1}\otimes\left(\mathbf{L}_{1t}^{-1}(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})\right)\right)\frac{\partial \mathrm{vec}\Big((\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})^\top\left(\mathbf{L}_{1t}^{-1}\right)^\top\Big)}{\partial \boldsymbol{\beta}_1^\top}\\ &\quad+\left(\left(\mathbf{L}_{1t}^{-1}(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})\right)\otimes \mathbf{I}_{d_1}\right)\frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t})\right)}{\partial \boldsymbol{\beta}_1^\top}\Bigg\}.\end{aligned}$$

On a aussi que pour $\mathbf{B}$ une matrice de dimension $m\times n$ (voir Harville (1997, p. 344)), $\mathrm{vec}\big(\mathbf{B}^\top\big)=\ \mathbf{K}_{mn}\mathrm{vec}(\mathbf{B})$. Ainsi :

$$\begin{aligned}
\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\beta}_1^\top} &= \mathbf{H}_{1d_1}\left\{\left(\mathbf{I}_{d_1}\otimes\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\right)\mathbf{K}_{d_1 1}+\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\otimes\mathbf{I}_{d_1}\right\}\times\\
&\quad \frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)}{\partial \boldsymbol{\beta}_1^\top},\\
&= \mathbf{H}_{1d_1}\left\{\left(\mathbf{I}_{d_1}\otimes\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\right)\mathbf{K}_{d_1 1}+\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\otimes\mathbf{I}_{d_1}\right\}\times\\
&\quad \frac{\partial \mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)}{\partial \boldsymbol{\beta}_1^\top},\\
&= \mathbf{H}_{1d_1}\left\{\left(\mathbf{I}_{d_1}\otimes\boldsymbol{\eta}_{1t}\right)\mathbf{K}_{d_1 1}+\boldsymbol{\eta}_{1t}\otimes\mathbf{I}_{d1}\right\}\frac{\partial \boldsymbol{\eta}_{1t}}{\partial \boldsymbol{\beta}_1^\top}.
\end{aligned}$$

De la relation précédente nous avons :

$$\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\beta}_1^\top} = n^{-1}\sum_{t=k+1}^{n}\mathbf{Q}_t\mathbf{P}_{t-k},$$

avec

$$\begin{aligned}
\mathbf{P}_{t-k} &= \mathbf{u}_{2,t-k}\otimes\mathbf{I}_{d_1'},\\
\mathbf{Q}_t &= \mathbf{H}_{1d_1}\left\{\left(\mathbf{I}_{d_1}\otimes\boldsymbol{\eta}_{1t}\right)\mathbf{K}_{d_1 1}+\boldsymbol{\eta}_{1t}\otimes\mathbf{I}_{d1}\right\}\frac{\partial \boldsymbol{\eta}_{1t}}{\partial \boldsymbol{\beta}_1^\top}.
\end{aligned}$$

L'hypothèse d'indépendance entre $\{\boldsymbol{\eta}_{1t}\}$ et $\{\boldsymbol{\eta}_{2t}\}$ ainsi que le lemme 2 dans l'article nous permet de conclure que :

$$\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\beta}_1^\top} = \mathbf{O}_P(n^{-1/2}).$$

<u>Deuxième cas :</u> $\boldsymbol{\alpha}=\boldsymbol{\omega}_1$

Des calculs similaires à ce qui a été fait auparavent donne :

$$\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\omega}_1^\top} = n^{-1}\sum_{t=k+1}^{n}\left(\mathbf{u}_{2,t-k}\otimes\mathbf{I}_{d_1'}\right)\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\omega}_1^\top}.$$

Comme dans le cas précédent nous obtenons aussi :

$$\begin{aligned}
\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\omega}_1^\top} &= \mathbf{H}_{1d_1}\left\{\left(\mathbf{I}_{d_1}\otimes\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\right)\mathbf{K}_{d_1 1}+\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\otimes\mathbf{I}_{d_1}\right\}\times\\
&\quad \frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)}{\partial \boldsymbol{\omega}_1^\top}.
\end{aligned}$$

Or nous vérifions que :

$$\begin{aligned}\frac{\partial \mathrm{vec}\big(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\big)}{\partial \boldsymbol{\omega}_1^\top} &= \mathbf{L}_{1t}^{-1}\frac{\partial\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)}{\partial \omega_1^\top} + \left(\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)^\top \otimes \mathbf{I}_{d_1}\right)\frac{\partial \mathrm{vec}\big(\mathbf{L}_{1t}^{-1}\big)}{\partial \boldsymbol{\omega}_1^\top},\\ &= \left(\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)^\top \otimes \mathbf{I}_{d_1}\right)\frac{\partial \mathrm{vec}\big(\mathbf{L}_{1t}^{-1}\big)}{\partial \boldsymbol{\omega}_1^\top},\\ &= -\left(\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\otimes \mathbf{I}_{d_1}\right)\left(\left(\mathbf{L}_{1t}^{-1}\right)^\top \otimes \mathbf{L}_{1t}^{-1}\right)\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top}.\end{aligned}$$

La dernière égalité provient de la propriété :

$$\frac{\partial \mathrm{vec}(\mathbf{F}^{-1})}{\partial \mathbf{X}^\top} = -\left(\left(\mathbf{F}^{-1}\right)^\top \otimes \mathbf{F}^{-1}\right)\frac{\partial \mathrm{vec}(\mathbf{F})}{\partial \mathbf{X}^\top}.$$

Voir en effet Harville (1997, p. 367). Calculons à présent $\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top}$. Rappelons que $\mathrm{vech}\big(\mathbf{L}_{1t}\mathbf{L}_{1t}^\top\big) = \mathbf{H}_{Ld_1}\mathrm{vec}\big(\mathbf{L}_{1t}\mathbf{L}_{1t}^\top\big)$. On a donc les égalités suivantes :

$$\begin{aligned}\frac{\partial \mathrm{vech}\big(\mathbf{L}_{1t}\mathbf{L}_{1t}^\top\big)}{\partial \boldsymbol{\omega}_1^\top} &= \mathbf{H}_{Ld_1}\left\{\left(\mathbf{I}_{d_1}\otimes \mathbf{L}_{1t}^\top\right)\frac{\partial \mathrm{vec}\mathbf{L}_{1t}^\top}{\partial \boldsymbol{\omega}_1^\top} + \left(\mathbf{L}_{1t}^\top \otimes \mathbf{I}_{d_1}\right)\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top}\right\},\\ &= \mathbf{H}_{Ld_1}\left\{\left(\mathbf{I}_{d_1}\otimes \mathbf{L}_{1t}^\top\right)\mathbf{K}_{d_1d_1}\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top} + \left(\mathbf{L}_{1t}^\top \otimes \mathbf{I}_{d_1}\right)\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top}\right\},\\ &= \mathbf{H}_{Ld_1}\left\{\left(\mathbf{I}_{d_1}\otimes \mathbf{L}_{1t}^\top\right)\mathbf{K}_{d_1d_1} + \left(\mathbf{L}_{1t}^\top \otimes \mathbf{I}_{d_1}\right)\right\}\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top},\\ &= \mathbf{H}_{Ld_1}\left(\mathbf{L}_{1t}^\top \otimes \mathbf{I}_{d_1}\right)\left\{\left(\mathbf{L}_{1t}^\top \otimes \mathbf{I}_{d_1}\right)^{-1}\left(\mathbf{I}_{d_1}\otimes \mathbf{L}_{1t}^\top\right)\mathbf{K}_{d_1d_1} + \mathbf{I}\right\}\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top}.\end{aligned}$$

Notons $\mathbf{G}_{Ld_1}$, l'inverse à gauche de $\mathbf{H}_{Ld_1}$. De plus comme $\mathbf{L}_{1t}\otimes \mathbf{I}_{d_1}$ est inversible, on peut écrire :

$$\left(\mathbf{L}_{1t}\otimes \mathbf{I}_{d_1}\right)^{-1}\mathbf{G}_{Ld_1}\frac{\partial \mathrm{vech}\big(\mathbf{L}_{1t}\mathbf{L}_{1t}^\top\big)}{\partial \boldsymbol{\omega}_1^\top} = \left(\mathbf{I} + \left(\mathbf{L}_{1t}^\top \otimes \mathbf{I}_{d_1}\right)^{-1}\left(\mathbf{I}_{d_1}\otimes \mathbf{L}_{1t}^\top\right)\mathbf{K}_{d_1d_1}\right)\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top}.$$

Posons $\mathbf{A} = -\left(\mathbf{L}_{1t}^\top \otimes \mathbf{I}_{d_1}\right)^{-1}\left(\mathbf{I}_{d_1}\otimes \mathbf{L}_{1t}^\top\right)\mathbf{K}_{d_1d_1}$ et supposons que $\|\mathbf{A}\| < 1$ ce qui entrainera que $(\mathbf{I}-\mathbf{A})$ est inversible. Ainsi :

$$\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top} = (\mathbf{I}-\mathbf{A})^{-1}\left(\mathbf{L}_{1t}^\top \otimes \mathbf{I}_{d_1}\right)^{-1}\mathbf{G}_{Ld_1}\frac{\partial \mathrm{vech}\big(\mathbf{L}_{1t}\mathbf{L}_{1t}^\top\big)}{\partial \boldsymbol{\omega}_1^\top}.$$

En utilisant l'hypothèse supplémentaire introduite sur $\mathbf{V}_{1t}$, on a que :

$$\frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}\mathbf{L}_{1t}^{\top}\right)}{\partial \boldsymbol{\omega}_1^{\top}} = \frac{\partial \mathrm{vech}(\mathbf{V}_{1t})}{\partial \boldsymbol{\omega}_1^{\top}} = \mathbf{I}_{d_1^2} - \sum_{j=1}^{\infty} \mathbf{A}_{1j}\left(\Lambda_1\right).$$

On a donc :

$$\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^{\top}} = (\mathbf{I}-\mathbf{A})^{-1}\left(\mathbf{L}_{1t}^{\top}\otimes \mathbf{I}_{d_1}\right)^{-1}\mathbf{G}_{Ld_1}\left(\mathbf{I}_{d_1^2} - \sum_{j=1}^{\infty} \mathbf{A}_{1j}\left(\Lambda_1\right)\right),$$

$$\begin{aligned}
\frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)}{\partial \boldsymbol{\omega}_1^{\top}} &= -\left(\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\otimes \mathbf{I}_{d_1}\right)\left(\left(\mathbf{L}_{1t}^{-1}\right)^{\top}\otimes \mathbf{L}_{1t}^{-1}\right)\times \\
&\quad (\mathbf{I}-\mathbf{A})^{-1}\left(\mathbf{L}_{1t}^{\top}\otimes \mathbf{I}_{d_1}\right)^{-1}\mathbf{G}_{Ld_1}\left(\mathbf{I}_{d_1^2} - \sum_{j=1}^{\infty} \mathbf{A}_{1j}\left(\Lambda_1\right)\right), \\
&= \mathbf{B},
\end{aligned}$$

et

$$\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\omega}_1^{\top}} = \mathbf{H}_{1d_1}\left\{\left(\mathbf{I}_{d_1}\otimes\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\right)\mathbf{K}_{d_1 1} + \left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\otimes \mathbf{I}_{d_1}\right\}\mathbf{B}.$$

On constate que l'expression précédente peut s'écrire :

$$\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\beta}_1^{\top}} = n^{-1}\sum_{t=k+1}^{n} \tilde{\mathbf{Q}}_t \tilde{\mathbf{P}}_{t-k},$$

où cette fois-ci :

$$\begin{aligned}
\tilde{\mathbf{P}}_{t-k} &= \mathbf{u}_{2,t-k}\otimes \mathbf{I}_{d_1'}, \\
\tilde{\mathbf{Q}}_t &= \frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\omega}_1^{\top}},
\end{aligned}$$

et où $\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\omega}_1^{\top}}$ est définit comme ci-dessus. Notons que dans le cas général en présumant que l'innovation a des moments d'ordre suffisamment élevé, l'argumentation précédente est valide.

Troisième cas : $\boldsymbol{\alpha} = \mathrm{vec}(\boldsymbol{\Lambda}_1)$

$$\begin{aligned}\frac{\partial \mathbf{u}_{1t}}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}} &= \mathbf{H}_{1d_1}\left\{\left(\mathbf{I}_{d_1}\otimes\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\right)\mathbf{K}_{d_1 1}+\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\otimes\mathbf{I}_{d_1}\right\}\times\\ &\quad \frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}}.\end{aligned}$$

De même que dans les cas précédents on a les égalités ci-dessous :

$$\begin{aligned}\frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}} &= \left(\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)^{\top}\otimes\mathbf{I}_{d_1}\right)\frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}},\\ &= \left(\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)^{\top}\otimes\mathbf{I}_{d_1}\right)\frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}},\\ &= -\left(\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\otimes\mathbf{I}_{d_1}\right)\left(\left(\mathbf{L}_{1t}^{-1}\right)^{\top}\otimes\mathbf{L}_{1t}^{-1}\right)\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}}.\end{aligned}$$

Or

$$\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}} = (\mathbf{I}-\mathbf{A})^{-1}\left(\left(\mathbf{L}_{1t}^{\top}\right)^{-1}\otimes\mathbf{I}_{d_1}\right)\mathbf{G}_{Ld_1}\frac{\partial \mathrm{vech}\left(\mathbf{L}_{1t}\mathbf{L}_{1t}^{\top}\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}},$$

et

$$\frac{\partial \mathrm{vech}\left(\mathbf{L}_{1t}\mathbf{L}_{1t}^{\top}\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}} = \sum_{j=1}^{\infty}\frac{\partial \mathbf{A}_{1j}\left(\boldsymbol{\Lambda}_1\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}}\left\{\mathrm{vech}\left(\boldsymbol{\epsilon}_{1,t-j}\boldsymbol{\epsilon}_{1,t-j}^{\top}\right)-\boldsymbol{\omega}_1\right\}.$$

De plus,

$$\frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}} = (\mathbf{I}-\mathbf{A})^{-1}\left(\left(\mathbf{L}_{1t}^{\top}\right)^{-1}\otimes\mathbf{I}_{d_1}\right)\mathbf{G}_{Ld_1}\sum_{j=1}^{\infty}\frac{\partial \mathbf{A}_{1j}\left(\boldsymbol{\Lambda}_1\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}}\left\{\mathrm{vech}\left(\boldsymbol{\epsilon}_{1,t-j}\boldsymbol{\epsilon}_{1,t-j}^{\top}\right)-\boldsymbol{\omega}_1\right\},$$

$$\begin{aligned}\frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}} &= -\left(\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\otimes\mathbf{I}_{d_1}\right)\left(\left(\mathbf{L}_{1t}^{-1}\right)^{\top}\otimes\mathbf{L}_{1t}^{-1}\right)\times\\ &\quad (\mathbf{I}-\mathbf{A})^{-1}\left(\left(\mathbf{L}_{1t}^{\top}\right)^{-1}\otimes\mathbf{I}_{d_1}\right)\mathbf{G}_{Ld_1}\sum_{j=1}^{\infty}\frac{\partial \mathbf{A}_{1j}\left(\boldsymbol{\Lambda}_1\right)}{\partial\left(\mathrm{vec}(\boldsymbol{\Lambda}_1)\right)^{\top}}\times\\ &\quad \left\{\mathrm{vech}\left(\boldsymbol{\epsilon}_{1,t-j}\boldsymbol{\epsilon}_{1,t-j}^{\top}\right)-\boldsymbol{\omega}_1\right\},\\ &= \mathbf{C}.\end{aligned}$$

Ainsi :

$$\frac{\partial \mathbf{u}_{1t}}{\partial\left(\text{vec}(\mathbf{\Lambda}_1)\right)^\top} \quad = \quad \mathbf{H}_{1d_1}\left\{\left(\mathbf{I}_{d_1}\otimes\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\right)\mathbf{K}_{d_1 1}+\left(\mathbf{L}_{1t}^{-1}\left(\mathbf{X}_{1t}-\boldsymbol{\mu}_{1t}\right)\right)\otimes\mathbf{I}_{d_1}\right\}\mathbf{C}.$$

On a enfin que $\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\beta}_1^\top}$ peut s'écrire sous la forme :

$$\frac{\partial \mathbf{c}_{\mathbf{Z}_1\mathbf{Z}_2}(k)}{\partial \boldsymbol{\beta}_1^\top} = n^{-1}\sum_{t=k+1}^{n}\mathbf{Q}_t^\star\mathbf{P}_{t-k}^\star,$$

où cependant nous avons les expressions :

$$\begin{aligned}\mathbf{P}_{t-k}^\star &= \mathbf{u}_{2,t-k}\otimes\mathbf{I}_{d_1'},\\ \mathbf{Q}_t^\star &= \frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\omega}_1^\top},\end{aligned}$$

et où $\frac{\partial \mathbf{u}_{1t}}{\partial \boldsymbol{\omega}_1^\top}$ est définit comme ci-dessus. Ceci conclut la preuve de la proposition.

3.2 Preuve de la proposition 2

Nous montrerons dans cette section que

$$\partial \mathbf{r}_{q_1q_2}/\partial \boldsymbol{\alpha}^\top = \mathbf{O}_P(n^{-1/2}),$$

où $\boldsymbol{\alpha}$ est l'un des coeffficients $\boldsymbol{\beta}_i$, $\boldsymbol{\omega}_i$, vec$(\mathbf{\Lambda}_i)$, $i = 1, 2$. Et partant du développement

$$\mathbf{r}_{\hat{q}_1\hat{q}_2} = \mathbf{r}_{q_1q_2} + \frac{\partial \mathbf{r}_{q_1q_2}}{\partial \boldsymbol{\alpha}^\top}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) + \mathbf{o}_P(n^{-1/2}),$$

nous déduirons que

$$\mathbf{r}_{\hat{q}_1\hat{q}_2} = \mathbf{r}_{q_1q_2} + \mathbf{o}_P(n^{-1/2}).$$

De ce résultat, il s'en suivra que la distribution asymptotique de $n^{1/2}\mathbf{r}_{\hat{q}_1\hat{q}_2}$ et $n^{1/2}\mathbf{r}_{q_1q_2}$ sont les mêmes. La démarche est similaire à celle de la Proposition 1 et est en fait beaucoup plus simple, compte tenu que la nature multivariée est substantiellement simplifiée. Pour tout $k \geq 0$ on va calculer
$\partial \mathbf{r}_{q_1q_2}(k)/\partial \boldsymbol{\alpha}^\top = \mathbf{O}_P(n^{-1/2})$. Pour des besoins de simplification on va poser :

$$\boldsymbol{\nu}_{it} = \frac{\mathbf{q}_{it} - \mathbf{d}_i}{\left(\mathbf{c}_{q_i q_i}(0)\right)^{1/2}} \quad i = 1, 2,$$

de sorte que :

$$\mathbf{r}_{\mathbf{q}_1\mathbf{q}_2}(k) = \begin{cases} n^{-1}\sum_{t=k+1}^{n} \boldsymbol{\nu}_{1t}\boldsymbol{\nu}_{2,t-k}, & k = 0, 1, ..., n-1, \\ \mathbf{r}_{q_2 q_1}(-k), & k = -n+1, ..., -1. \end{cases}$$

De façon générique, en dérivant par rapport à un vecteur $\boldsymbol{\alpha}$ générique nous obtenons :

$$\begin{aligned} \frac{\partial \mathbf{r}_{q_1 q_2}}{\partial \boldsymbol{\alpha}^\top}(k) &= \frac{\partial\left(n^{-1}\sum_{t=k+1}^{n} \boldsymbol{\nu}_{1t}\boldsymbol{\nu}_{2,t-k}\right)}{\partial \boldsymbol{\alpha}^\top}, \\ &= n^{-1}\sum_{t=k+1}^{n}\left(\frac{\partial \boldsymbol{\nu}_{1t}}{\partial \boldsymbol{\alpha}^\top}\boldsymbol{\nu}_{2,t-k} + \boldsymbol{\nu}_{1t}\frac{\partial \boldsymbol{\nu}_{2,t-k}}{\partial \boldsymbol{\alpha}^\top}\right). \end{aligned}$$

<u>Premier cas :</u> $\boldsymbol{\alpha} = \boldsymbol{\beta}_1$

En dérivant par rapport à $\boldsymbol{\beta}_1$, nous avons :

$$\frac{\partial \mathbf{r}_{q_1 q_2}}{\partial \boldsymbol{\beta}_1^\top}(k) = n^{-1}\sum_{t=k+1}^{n}\frac{\partial \boldsymbol{\nu}_{1t}}{\partial \boldsymbol{\beta}_1^\top}\boldsymbol{\nu}_{2,t-k},$$

où

$$\begin{aligned} \frac{\partial \boldsymbol{\nu}_{1t}}{\partial \boldsymbol{\beta}_1^\top} &= \left(\mathbf{c}_{q_1 q_1}(0)\right)^{-1/2}\frac{\partial\left(\boldsymbol{\epsilon}_{1t}^\top \mathbf{V}_{1t}^{-1}\boldsymbol{\epsilon}_{1t}\right)}{\partial \boldsymbol{\beta}_1^\top}, \\ &= 2\left(\mathbf{c}_{q_1 q_1}(0)\right)^{-1/2}\frac{\partial \boldsymbol{\epsilon}_{1t}^\top}{\partial \boldsymbol{\beta}_1^\top}\mathbf{V}_{1t}^{-1}\boldsymbol{\epsilon}_{1t}, \\ &= -2\left(\mathbf{c}_{q_1 q_1}(0)\right)^{-1/2}\frac{\partial \boldsymbol{\mu}_{1t}^\top}{\partial \boldsymbol{\beta}_1^\top}\mathbf{V}_{1t}^{-1}\boldsymbol{\epsilon}_{1t}. \end{aligned}$$

Enfin on a :

$$\frac{\partial \mathbf{r}_{q_1 q_2}}{\partial \boldsymbol{\beta}_1^\top}(k) = -2n^{-1}\left(\mathbf{c}_{q_1 q_1}(0)\right)^{-1/2}\sum_{t=k+1}^{n}\left(\frac{\partial \boldsymbol{\mu}_{1t}^\top}{\partial \boldsymbol{\beta}_1^\top}\mathbf{V}_{1t}^{-1}\boldsymbol{\epsilon}_{1t}\right)\boldsymbol{\nu}_{2,t-k}.$$

<u>Deuxième cas :</u> $\boldsymbol{\alpha} = \boldsymbol{\omega}_1$

$$\frac{\partial \mathbf{r}_{q_1 q_2}}{\partial \boldsymbol{\omega}_1^\top}(k) = n^{-1}\sum_{t=k+1}^{n}\frac{\partial \boldsymbol{\nu}_{1t}}{\partial \boldsymbol{\omega}_1^\top}\boldsymbol{\nu}_{2,t-k},$$

$$\frac{\partial \boldsymbol{\nu}_{1t}}{\partial \omega_1^\top} = \left(\mathbf{c}_{q_1q_1}(0)\right)^{-1/2} \frac{\partial \left(\boldsymbol{\epsilon}_{1t}^\top \mathbf{V}_{1t}^{-1} \boldsymbol{\epsilon}_{1t}\right)}{\partial \boldsymbol{\omega}_1^\top}.$$

Or $\mathbf{V}_{1t} = \mathbf{L}_{1t}\mathbf{L}_{1t}^\top$ ce qui entraîne :

$$\begin{aligned}
\boldsymbol{\epsilon}_{1t}^\top \mathbf{V}_{1t} \boldsymbol{\epsilon}_{1t} &= \boldsymbol{\epsilon}_{1t}^\top \left(\mathbf{L}_{1t}^\top\right)^{-1} \mathbf{L}_{1t}^{-1} \boldsymbol{\epsilon}_{1t}, \\
&= \left(\mathbf{L}_{1t}^{-1}\boldsymbol{\epsilon}_{1t}\right)^\top \left(\mathbf{L}_{1t}\boldsymbol{\epsilon}_{1t}\right), \\
&= \left\{\mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\boldsymbol{\epsilon}_{1t}\right)\right\}^\top \left\{\mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\boldsymbol{\epsilon}_{1t}\right)\right\}, \qquad (3.1)\\
&= \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)^\top \left(\boldsymbol{\epsilon}_{1t} \otimes \mathbf{I}_{d_1}\right) \left(\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right), \qquad (3.2)\\
&= \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)^\top \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right).
\end{aligned}$$

Les égalités (3.1) et (3.2) proviennent respectivement des relations $\mathbf{X}^\top\mathbf{X} = (\mathrm{vec}(\mathbf{X}))^\top \mathrm{vec}(\mathbf{X})$ et $\mathrm{vec}(\mathbf{ABC}) = \left(\mathbf{C}^\top \otimes \mathbf{A}\right) \mathrm{vec}\mathbf{B}$. On a donc :

$$\begin{aligned}
\left(\mathbf{c}_{q_1q_1}(0)\right)^{1/2} \frac{\partial \boldsymbol{\nu}_{1t}}{\partial \boldsymbol{\omega}_1^\top} &= \frac{\partial \left(\boldsymbol{\epsilon}_{1t}^\top \mathbf{V}_{1t}^{-1} \boldsymbol{\epsilon}_{1t}\right)}{\partial \boldsymbol{\omega}_1^\top}, \\
&= \frac{\partial \left(\mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)^\top \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)\right)}{\partial \boldsymbol{\omega}_1^\top}, \\
&= \left(\frac{\partial \left(\mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)\right)^\top}{\partial \boldsymbol{\omega}_1^\top}\right) \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right) \\
&+ \left(\mathrm{vec}(\mathbf{L}_{1t})^{-1}\right)^\top \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)}{\partial \boldsymbol{\omega}_1^\top}, \\
&= \left(\frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)}{\partial \boldsymbol{\omega}_1^\top}\right)^\top \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right) \\
&+ \left(\mathrm{vec}(\mathbf{L}_{1t})^{-1}\right)^\top \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)}{\partial \boldsymbol{\omega}_1^\top}, \\
&= \left(-\left(\left(\mathbf{L}_{1t}^{-1}\right)^\top \otimes \mathbf{L}_{1t}^{-1}\right) \frac{\partial \mathrm{vec}(\mathbf{L}_{1t})}{\partial \boldsymbol{\omega}_1^\top}\right)^\top \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right) \\
&- \left(\mathrm{vec}(\mathbf{L}_{1t})^{-1}\right)^\top \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \left(\left(\mathbf{L}_{1t}^{-1}\right)^\top \otimes \mathbf{L}_{1t}^{-1}\right) \frac{\partial \mathrm{vec}\left(\mathbf{L}_{1t}^{-1}\right)}{\partial \boldsymbol{\omega}_1^\top}.
\end{aligned}$$

On avait établi dans la preuve de la proposition 1 que sous certaines conditions,

$$\frac{\partial \mathrm{vec}\mathbf{L}_{1t}}{\partial \boldsymbol{\omega}_1^\top} = (\mathbf{I} - \mathbf{A})^{-1} \left(\mathbf{L}_{1t}^\top \otimes \mathbf{I}_{d_1}\right)^{-1} \mathbf{G}_{Ld_1} \left(\mathbf{I}_{d_1^2} - \sum_{j=1}^{\infty} \mathbf{A}_{1j}(\Lambda_1)\right).$$

En remplaçant la valeur de $\frac{\partial \text{vec}\mathbf{L}_{1t}}{\partial \boldsymbol{\omega}_1^\top}$ dans la dernière égalité ci-dessus on a une expression de $(\mathbf{c}_{q_1q_1}(0))^{1/2} \frac{\partial \boldsymbol{\nu}_{1t}}{\partial \boldsymbol{\omega}_1^\top}$ et on conclut comme dans les cas précédents.

<u>Troisième cas :</u> $\boldsymbol{\alpha} = \text{vec}(\boldsymbol{\Lambda}_1)$

On débute en calculant :

$$\frac{\partial \mathbf{r}_{q_1q_2}}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} = n^{-1} \sum_{t=k+1}^{n} \frac{\partial \boldsymbol{\nu}_{1t}}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} \boldsymbol{\nu}_{2,t-k},$$

et

$$\begin{aligned}
\frac{\partial \boldsymbol{\nu}_{1t}}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} &= \frac{\partial}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} \left\{ \text{vec}\big(\mathbf{L}_{1t}^{-1}\big)^\top \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \text{vec}\big(\mathbf{L}_{1t}^{-1}\big) \right\}, \\
&= 2 \left(\frac{\partial \left(\text{vec}\big(\mathbf{L}_{1t}^{-1}\big)\right)^\top}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} \right) \left(\boldsymbol{\epsilon}_{1t}\boldsymbol{\epsilon}_{1t}^\top \otimes \mathbf{I}_{d_1}\right) \text{vec}\big(\mathbf{L}_{1t}^{-1}\big).
\end{aligned}$$

Calculons à présent $\frac{\partial \left\{\text{vec}\left(\mathbf{L}_{1t}^{-1}\right)\right\}^\top}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top}$.

On a :

$$\begin{aligned}
\frac{\partial \left\{\text{vec}\big(\mathbf{L}_{1t}^{-1}\big)\right\}^\top}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} &= \mathbf{K}_{d_1^2}^\top \frac{\partial \left(\text{vec}\big(\mathbf{L}_{1t}^{-1}\big)\right)}{\partial \left(\text{vec}(\boldsymbol{\Lambda}_1)\right)^\top}, \\
&= -\mathbf{K}_{d_1^2}^\top \left\{ \big(\mathbf{L}_{1t}^{-1}\big)^\top \otimes \mathbf{L}_{1t}^{-1} \right\} \frac{\partial \text{vec}(\mathbf{L}_{1t})}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top}.
\end{aligned}$$

Dans un calcul antérieur nous avions obtenu :

$$\begin{aligned}
\frac{\partial \text{vec}(\mathbf{L}_{1t})}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} &= (\mathbf{I} - \mathbf{A})^{-1} \left(\big(\mathbf{L}_{1t}^\top\big)^{-1} \otimes \mathbf{I}_{d_1} \right) \mathbf{G}_{Ld_1} \sum_{j=1}^{\infty} \frac{\partial}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} \mathbf{A}_{1j}(\boldsymbol{\Lambda}_1) \times \\
&\quad \left\{ \text{vech}\big(\boldsymbol{\epsilon}_{1,t-j}\boldsymbol{\epsilon}_{1,t-j}^\top\big) - \boldsymbol{\omega}_1 \right\}.
\end{aligned}$$

Ainsi, nous avons :

$$\begin{aligned}
\frac{\partial \left\{\text{vec}\big(\mathbf{L}_{1t}^{-1}\big)\right\}^\top}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} &= -\mathbf{K}_{d_1^2}^\top \left\{ \big(\mathbf{L}_{1t}^{-1}\big)^\top \otimes \mathbf{L}_{1t}^{-1} \right\} (\mathbf{I} - \mathbf{A})^{-1} \left(\big(\mathbf{L}_{1t}^\top\big)^{-1} \otimes \mathbf{I}_{d_1} \right) \times \\
&\quad \mathbf{G}_{Ld_1} \sum_{j=1}^{\infty} \frac{\partial}{\partial \left\{\text{vec}(\boldsymbol{\Lambda}_1)\right\}^\top} \mathbf{A}_{1j}(\Lambda_1) \left\{ \text{vech}\big(\boldsymbol{\epsilon}_{1,t-j}\boldsymbol{\epsilon}_{1,t-j}^\top\big) - \boldsymbol{\omega}_1 \right\}.
\end{aligned}$$

En remplaçant la valeur de $\frac{\partial\{\text{vec}(\mathbf{L}_{1t}^{-1})\}^\top}{\partial\{\text{vec}(\mathbf{\Lambda}_1)\}^\top}$ dans l'expression de $\frac{\partial\boldsymbol{\nu}_{1t}}{\partial\{\text{vec}(\mathbf{\Lambda}_1)\}^\top}$ et le résultat obtenu dans celui de $\frac{\partial\mathbf{r}_{q_1q_2}}{\partial\{\text{vec}(\mathbf{\Lambda}_1)\}^\top}$, on obtient une expression qui nous permet de conclure comme dans les cas précédents.

Chapitre 4

Simulations complémentaires

Dans cette partie, nous présentons quelques résultats de simulation complémentaires, qui n'ont pas été retenu dans l'article. Nous allons considérer les mêmes processus générateurs de données (DGP) que dans l'article mais avec des paramètres différents. Nous nous limiterons ici à une étude de puissance pour les tests $\mathbf{ER}_{nM}$, $\mathbf{LL}_{nM}$, pour $n = 1, \ldots, 10$ et $\mathbf{LL}_n(k)$, $\mathbf{ER}_n(k)$ pour $k = -10, \ldots, 10$, ainsi que leurs versions modifiées.

4.1 Description des expériences

Afin de compléter l'étude des propriétés en échantillons finis, nous allons considérer les DGP suivants :

$$\mathbf{X}_{it} = \mathbf{L}_{it}\boldsymbol{\eta}_{it} \; i = 1, 2, \tag{4.1}$$

$$\mathbf{V}_{it} = \mathbf{L}_{it}\mathbf{L}_{it}^{\top}, \tag{4.2}$$

$$\begin{aligned} \mathbf{V}_{1t} = {} & \boldsymbol{\Theta} + \mathbf{A}_1^{\top}\left(\boldsymbol{\epsilon}_{1,t-1}\boldsymbol{\epsilon}_{1,t-1}^{\top}\right)\mathbf{A}_1 + \mathbf{B}_1^{\top}\mathbf{V}_{1,t-1}\mathbf{B}_1 + \\ & \delta_{12}\mathbf{C}_1^{\top}\left(\boldsymbol{\epsilon}_{2,t-l_1}\boldsymbol{\epsilon}_{2,t-l_1}^{\top}\right)\mathbf{C}_1 + \gamma_{12}\mathbf{D}_1^{\top}\mathbf{V}_{2,t-l_1}\mathbf{D}_1, \end{aligned} \tag{4.3}$$

$$\begin{aligned} \mathbf{V}_{2t} = {} & \boldsymbol{\Theta} + \mathbf{A}_2^{\top}\boldsymbol{\epsilon}_{2,t-1}\boldsymbol{\epsilon}_{2,t-1}^{\top}\mathbf{A}_2 + \mathbf{B}_2^{\top}\mathbf{V}_{2,t-1}\mathbf{B}_2 + \\ & \delta_{21}\mathbf{C}_2^{\top}\left(\boldsymbol{\epsilon}_{1,t-l_2}\boldsymbol{\epsilon}_{2,t-l_2}^{\top}\right)\mathbf{C}_2 + \gamma_{21}\mathbf{D}_2^{\top}\mathbf{V}_{1,t-l_2}\mathbf{D}_2. \end{aligned} \tag{4.4}$$

Les valeurs $\delta_{12} = \gamma_{12} = \delta_{21} = \gamma_{21} = 0$ correspondent à l'hypothèse nulle et sont utiles pour l'étude de niveau. Mais pour l'étude de puissance nous ferons varier ces coefficients afin d'introduire une relation de causalité en variance. Dans un premier modèle S_1, nous supposerons $\delta_{12} = \gamma_{12} = \delta_{21} = \gamma_{21} = 0.49$ ce qui induit une causalité bidirectionnelle ; nous prendrons aussi $\mathbf{A}_1 = \mathbf{A}_2 = \mathbf{\Phi_1}$, $\mathbf{B}_1 = \mathbf{\Phi}_2$, $\mathbf{B}_2 = \mathbf{\Phi}_3$, $\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{\Phi}_4$, $\mathbf{D}_1 = \mathbf{D}_2 = \mathbf{\Phi}_3$ et enfin $l_1 = l_2 = 2$. Cette situation correspond à une causalité bidirectionnelle. Dans un deuxième modèle S_2, nous utiliserons le même processus générateur mais avec les valeurs $\delta_{12} = \gamma_{12} = \delta_{21} = \gamma_{21} = 0.10$.

4.2 Discussion des résultats

Dans les tableaux 4.1 et 4.2, on observe les résultats du modèle S_1. Sans surprise, les tests de délai global suggèrent un lien de causalité entre les deux séries. Par contre dans le tableau 4.2, contrairement à ce qu'on se serait attendu, la causalité est unidirectionnelle. Les quatre tests sont significatifs pour les délais positifs et non pour les délais négatifs ; la série $\mathbf{X}_{2t}$ cause la série $\mathbf{X}_{1t}$ et non l'inverse. Le choix des valeurs $\delta_{i,j}$ et $\gamma_{i,j}$, $i, j = 1, 2$ ainsi que les matrices $\mathbf{A}_i$, $\mathbf{B}_i$ et $\mathbf{C}_i$, $i = 1, 2$ semblent être à l'origine de cette situation. Le modèle S_2 a d'ailleurs été étudié en détail pour plus d'éclaircissement. En choisissant les valeurs $\delta_{12} = \gamma_{12} = \delta_{21} = \gamma_{21} = 0.10$ pour ce modèle, on commence à apercevoir la causalité dans l'autre sens notamment aux délais -3 et -4. Nous avons expérimenté plusieurs variantes de ce modèle (en faisant varier les paramètres ainsi que les tailles d'échantillon) et nous avons abouti à chaque fois à une causalité dans les deux directions. Les résultats complets n'ont pas été reporté dans le présent document.

Tableau 4.1 – Puissance empirique, modèle S_1, niveaux nominaux 5% et 10%, $n = 1000, 2000$, pour les tests globaux $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ et $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
M	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
1	5.5	5.6	3.1	3.1	4.3	4.3	4.1	4.1	9.9	9.9	8.8	8.8	8.9	9.0	9.3	9.3
2	12.6	12.6	11.2	11.3	19.2	19.2	20.3	20.3	21.0	21.5	19.0	19.0	28.2	28.2	28.4	28.4
3	23.4	24.1	23.3	23.3	44.7	44.7	43.4	43.5	33.9	34.2	31.9	32.1	57.4	57.8	53.5	53.6
4	34.4	34.6	31.6	31.9	62.1	62.2	56.5	56.6	45.9	46.3	39.5	39.7	72.4	73.1	67.6	67.7
5	38.3	39.2	35.1	35.1	71.4	71.5	63.4	63.5	52.6	53.0	45.4	45.6	79.1	79.4	73.6	73.9
6	41.2	43.1	37.4	37.8	75.8	76.5	69.5	69.8	55.0	55.6	49.0	49.3	83.3	83.6	77.9	77.9
7	45.0	45.6	37.8	38.2	78.3	78.6	70.1	70.3	56.1	56.9	49.7	50.0	85.5	85.9	79.6	79.6
8	44.9	45.9	38.6	39.1	78.5	78.8	70.4	71.1	56.6	57.6	49.9	50.3	87.0	87.2	80.1	80.2
9	44.8	46.4	37.9	38.3	78.9	79.6	71.3	72.0	57.5	58.8	49.1	49.8	86.3	86.7	80.1	80.5
10	45.5	47.7	37.9	38.3	78.8	79.3	71.1	71.1	57.9	59.9	48.0	48.8	86.9	87.4	81.1	81.6

Tableau 4.2 – Puissance empirique, modèle S_1, niveaux nominaux 5% et 10%, $n = 1000, 2000$, pour les tests aux délais individuels $\mathrm{ER}_n(k)$, $\mathrm{ER}_n^*(k)$, $\mathrm{LL}_n(k)$ et $\mathrm{LL}_n^*(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
k	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
-10	6.5	6.9	5.2	5.4	6.5	6.5	5.5	5.5	11.6	11.8	10.5	10.6	11.2	11.5	11.0	11.0
-9	6.1	6.4	5.1	5.1	5.7	5.7	4.9	4.9	10.0	10.3	8.3	8.4	10.4	10.6	8.6	8.7
-8	5.0	5.1	5.3	5.3	5.6	5.6	4.8	4.8	9.3	9.4	10.1	10.2	9.3	9.5	9.9	9.9
-7	5.2	5.6	4.6	4.7	6.0	6.0	5.1	5.2	10.7	11.0	8.9	8.9	9.8	9.8	11.0	11.1
-6	6.4	6.4	4.2	4.2	5.1	5.3	6.0	6.0	10.7	11.2	9.7	9.8	10.0	10.1	9.8	9.9
-5	4.9	5.0	4.8	4.8	4.8	4.8	5.4	5.5	9.0	9.1	9.2	9.2	8.2	8.4	10.7	10.8
-4	5.3	5.4	4.6	4.6	6.5	6.6	4.2	4.2	10.9	11.0	9.4	9.4	11.6	11.6	9.1	9.1
-3	6.6	6.7	6.4	6.4	5.9	5.9	6.3	6.3	12.0	12.4	11.4	11.4	11.0	11.1	13.5	13.5
-2	6.1	6.2	5.2	5.2	5.9	5.9	5.7	5.7	10.2	10.5	9.8	9.8	9.4	9.5	11.9	11.9
-1	4.8	4.8	4.4	4.5	4.9	4.9	5.0	5.0	9.2	9.2	9.5	9.5	9.3	9.3	11.0	11.0
0	5.6	5.6	3.7	3.7	4.9	4.9	4.8	4.8	9.8	9.8	8.7	8.7	9.3	9.3	8.8	8.8
1	5.8	5.8	5.0	5.0	5.9	5.9	5.3	5.3	10.1	10.1	10.5	10.5	11.1	11.1	10.3	10.3
2	24.2	24.2	19.3	19.4	43.3	43.3	35.4	35.6	33.9	33.9	29.6	29.6	54.4	54.5	48.0	48.0
3	33.5	33.6	28.3	28.3	59.3	59.6	50.9	51.0	45.0	45.2	40.0	40.1	70.1	70.2	63.9	64.0
4	33.7	33.8	25.3	25.3	53.5	53.6	44.9	44.9	43.8	44.1	37.1	37.1	63.8	63.8	55.5	55.7
5	24.1	24.4	19.1	19.4	41.7	41.9	34.1	34.1	33.1	33.5	28.2	28.3	53.3	53.4	45.3	45.3
6	20.1	20.2	17.8	18.1	32.3	32.3	28.4	28.4	28.6	28.9	25.9	26.1	42.1	42.6	39.0	39.1
7	17.8	18.1	13.7	13.9	26.2	26.3	23.1	23.1	25.8	26.2	20.6	20.6	36.3	36.8	33.8	33.8
8	15.5	16.3	11.6	11.8	20.8	21.1	17.1	17.3	22.8	23.1	20.2	20.7	29.3	29.3	27.1	27.2
9	12.5	13.0	9.9	10.0	18.8	18.9	16.1	16.1	19.2	19.6	15.3	15.6	27.1	27.4	25.5	25.5
10	11.3	11.9	7.6	7.7	16.7	17.1	13.5	13.5	17.7	18.4	12.9	13.2	24.6	24.9	20.9	21.2

Tableau 4.3 – Puissance empirique, modèle S_2, niveaux nominaux 5% et 10%, $n = 1000, 2000$, pour les tests globaux $\mathrm{ER}_{n,M}$, $\mathrm{ER}^*_{n,M}$, $\mathrm{LL}_{n,M}$ et $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
M	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
1	6.1	6.2	3.1	3.1	4.4	4.4	4.5	4.5	9.9	9.9	8.1	8.1	8.4	8.4	8.9	8.9
2	27.0	27.2	23.5	23.8	50.3	50.5	45.9	45.9	37.4	37.8	33.8	33.9	63.2	63.6	57.7	57.8
3	58.4	58.9	48.4	48.7	91.7	92.0	82.7	82.8	68.5	68.6	60.8	60.8	94.9	94.9	88.9	88.9
4	76.3	76.5	63.7	64.0	99.6	98.7	94.4	94.4	84.6	84.7	74.6	74.6	99.6	99.6	96.9	96.9
5	82.9	83.1	70.7	70.9	99.7	99.6	97.2	97.2	89.8	90.0	79.7	79.8	99.8	99.8	98.6	98.6
6	86.4	87.1	74.3	74.7	99.8	99.7	98.0	98.0	91.5	92.2	83.4	83.7	99.9	99.9	98.9	98.9
7	87.0	87.6	74.9	75.0	99.8	99.8	98.2	98.2	92.4	92.9	84.6	85.0	100.0	100.0	99.0	99.0
8	88.0	89.0	76.9	77.2	99.9	99.8	98.5	98.5	93.1	93.5	85.0	85.2	100.0	100.0	99.2	99.2
9	88.2	89.0	76.8	77.0	99.9	99.9	98.6	98.6	93.5	94.0	85.0	85.5	99.9	99.9	99.4	99.4
10	87.8	88.4	76.4	76.7	99.8	99.8	98.8	98.8	92.5	93.2	85.1	85.6	99.9	99.9	99.3	99.3

Tableau 4.4 – Puissance empirique, modèle S_2, niveaux nominaux 5% and 10%, $n = 1000, 2000$, pour les tests aux délais individuels $\mathrm{ER}_n(k)$, $\mathrm{ER}_n^*(k)$, $\mathrm{LL}_n(k)$ et $\mathrm{LL}_n^*(k)$, $k = -10, \ldots, -1, 0, 1, \ldots, 10$.

	$\alpha = 5\%$								$\alpha = 10\%$							
	$n = 1000$				$n = 2000$				$n = 1000$				$n = 2000$			
k	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*	ER	ER*	LL	LL*
-10	6.8	7.0	5.3	5.4	6.4	6.4	5.7	5.7	12.2	12.7	10.3	10.5	12.0	12.6	11.1	11.2
-9	6.2	6.5	4.7	4.8	5.8	5.8	5.4	5.4	10.3	10.7	8.5	8.5	10.7	10.9	9.5	9.7
-8	5.2	5.3	4.9	5.2	5.9	5.9	4.9	4.9	8.5	9.0	9.9	10.3	10.6	10.8	10.8	10.9
-7	6.5	6.7	4.9	4.9	5.8	5.8	5.7	5.7	11.3	11.5	9.0	9.1	10.7	10.9	10.9	10.9
-6	6.4	6.4	4.7	4.7	6.5	6.6	6.6	6.6	11.1	11.4	10.3	10.3	11.2	11.3	11.2	11.2
-5	5.8	5.9	5.3	5.3	6.4	6.5	6.5	6.5	9.9	10.0	9.7	9.9	9.9	10.1	11.9	12.0
-4	6.5	6.5	5.9	6.0	8.3	8.4	6.4	6.4	12.5	12.8	10.1	10.2	14.4	14.6	11.2	11.2
-3	8.7	8.8	7.8	7.8	8.8	8.9	9.0	9.0	15.4	14.5	13.0	13.1	14.9	11.0	15.6	15.6
-2	6.8	6.8	5.3	5.3	6.2	6.2	5.8	5.8	11.9	12.0	11.0	11.0	11.2	11.2	11.7	11.7
-1	4.6	4.6	4.4	4.4	5.1	5.1	5.2	5.2	9.1	9.1	10.5	10.5	9.4	9.4	10.9	10.9
0	5.8	5.8	3.6	3.6	3.8	3.8	4.3	4.3	9.4	9.4	9.1	9.1	9.9	9.9	8.5	8.5
1	5.2	5.3	4.7	4.7	5.2	5.2	4.7	4.7	9.4	9.4	11.1	11.1	10.4	10.4	9.8	9.9
2	54.7	55.0	40.5	40.5	81.4	81.6	68.7	68.8	63.6	63.7	53.1	53.2	89.0	89.0	77.8	77.8
3	57.7	68.8	55.3	55.3	94.9	94.9	85.4	85.4	76.6	76.9	65.4	65.4	97.4	97.4	91.4	91.5
4	61.3	61.6	49.3	49.5	88.9	88.9	75.9	76.0	70.4	70.5	59.4	59.4	92.7	92.7	83.5	83.5
5	44.3	44.8	33.5	33.6	75.2	75.4	58.7	58.9	55.7	55.8	45.5	45.6	81.8	81.8	69.3	69.3
6	32.9	33.1	28.8	29.0	57.3	57.4	47.4	47.5	43.7	44.1	37.8	38.2	68.0	68.1	58.8	58.8
7	29.0	29.0	21.0	21.0	47.6	48.2	40.1	40.1	38.5	39.0	28.9	29.1	60.5	61.0	52.4	52.4
8	23.7	23.8	17.4	17.6	36.7	37.2	30.6	30.7	31.1	31.4	27.5	27.6	48.6	48.7	41.8	42.1
9	17.9	18.7	13.4	13.5	32.6	32.9	26.4	26.5	26.6	27.3	20.4	20.5	43.3	43.5	36.6	36.8
10	15.0	15.3	11.1	11.6	26.4	26.5	20.1	20.3	22.4	23.0	18.3	18.4	36.6	36.6	30.0	30.0

Chapitre 5

Étude de cas supplémentaire

Dans l'article, afin d'illustrer comment nos tests peuvent servir à détecter la causalité en variance avec des données réelles, nous avons étudié deux séries nord-américaines que nous avons comparé à deux séries européennes. Dans cette partie, nous ferons une étude similaire. Nous allons considérer de nouveau les deux séries européennes que nous allons comparer cette fois à deux séries provenant du marché asiatique. L'indice boursier Hang Seng qui est l'indice phare du marché boursier de Hong Kong et qui regroupe 43 actions de sociétés qui sont cotées sur la Bourse de Hong Kong. Par la suite l'indice Nikkei225 qui est le principal indice boursier de la Bourse japonaise, composé des 225 premières capitalisation japonaise. Le rapport ci-dessous est assez sommaire compte tenue de la similarité de la démarche avec l'étude menée dans l'article.

5.1 Description des indices et notations

La première série est constituée de deux indices européens à savoir l'indice du marché suisse ainsi que l'indice de la place de Paris, le CAC40. Tout comme dans l'article, les données sont journalières (du lundi au vendredi) allant du 4 janvier 2007 au 30 décembre 2009. Après exclusion des données manquantes dû pour la plupart aux jours fériés, nous avons retenu 699 observations au total dans l'analyse. Les séries résultantes sont notées $\{SMI_t, t = 1, \cdots, n\}$ et $\{CAC40_t, t = 1, \cdots, n\}$ en ce qui concerne les indices européens, puis $\{N225_t, t = 1, \cdots, n\}$ et $\{HIS_t, t = 1, \cdots, n\}$ pour les indices japonais et chinois respectivement.

La figure 5.1 montre les séries brutes. On peut remarquer qu'elles ne sont pas station-

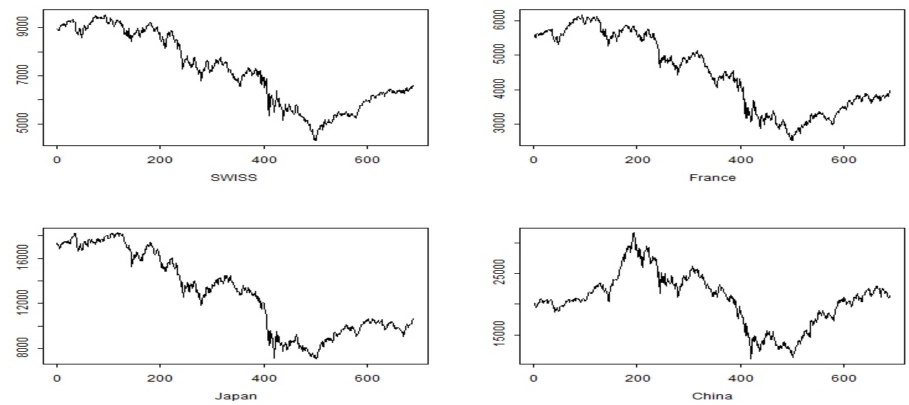


FIGURE 5.1 – Indice du marché Suisse, indice CAC 40, indice N225 et indice HIS de janvier 2007 à décembre 2009.

.

naires. Nous avons appliqué la transformation logarithmique puis le filtre $1 - B$ sur les données originales afin de calculer la série des rendements. La figure 5.2 représente les séries résultantes qui contrairement aux précédentes ne sont pas incompatibles avec la stationnarité.

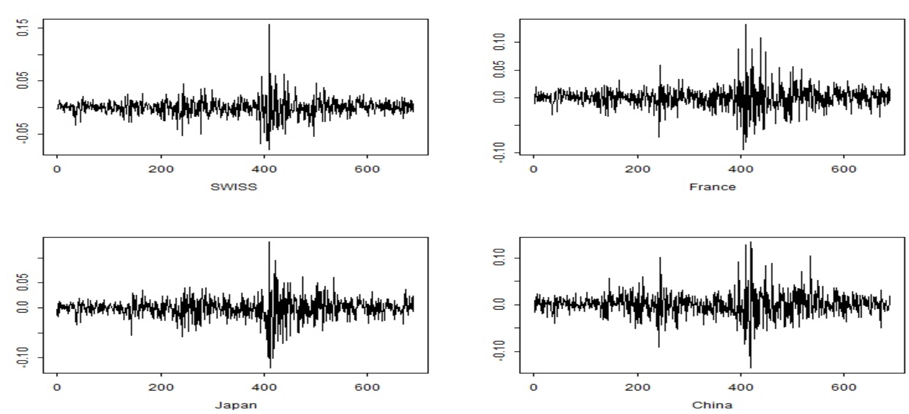


FIGURE 5.2 – Séries des rendements.

Les séries utilisées pour l'analyse sont donc :

$$\mathbf{X}_{1t} = \begin{pmatrix} X_{1t}(1) \\ X_{1t}(2) \end{pmatrix} = \begin{pmatrix} (1-B)\log(\mathrm{SMI}_t) \\ (1-B)\log(\mathrm{CAC40}_t) \end{pmatrix},$$
$$\mathbf{X}_{2t} = \begin{pmatrix} X_{2t}(1) \\ X_{2t}(2) \end{pmatrix} = \begin{pmatrix} (1-B)\log(\mathrm{N225}_t) \\ (1-B)\log(\mathrm{HIS}_t) \end{pmatrix}.$$

Dans la prochaine section, nous décrivons les modèles utilisés lors de l'estimation et appliquons les tests de causalité en variance.

5.2 Test de causalité en variance

La méthodologie d'analyse est la même que celle utilisée dans l'article. Nous n'allons pas proposer un modèle global pour toutes les séries. Nous modélisons chaque série vectorielle séparément puis nous effectuons les tests de causalité en variance. Après exploration de plusieurs modèles de type VARMA-GARCH, nous avons finalement retenu un bruit blanc faible pour les deux séries. Les modèles sont donc les suivants :

$$\begin{aligned} \mathbf{X}_{1t} &= \boldsymbol{\epsilon}_{1t}, \\ \mathbf{X}_{2t} &= \boldsymbol{\epsilon}_{2t}, \end{aligned}$$

et pour chacun de ces modèles la variance conditionnelle est supposée de la forme :

$$\mathbf{V}_t = \mathbf{C} + \mathbf{A} \odot (\mathbf{X}_{1,t-1}\mathbf{X}_{1,t-1}^{\top}) + \mathbf{B} \odot \mathbf{V}_{t-1}.$$

Les estimations des coefficients de ces modèles sont données dans le tableau 5.1. Ils sont tous significativement différents de zéro au niveau 5%. Nous avons aussi produit dans le tableau 5.2 les statistiques de Hosking $\mathbf{H}_n^{(i,i)}(k)$ et $\mathbf{HM}_n^{(i,i)}(k)$, $i = 1, 2$ et $k = 1, \ldots, 12$ sur les résidus standardisés. On peut remarquer que tout comme le précédent le bruit blanc est assez satisfaisant comme modèle pour $\mathbf{X}_{1,t}$ et $\mathbf{X}_{2,t}$ (se référer à l'article pour les détails de ce test). Notons tout de même que ce résultat n'est pas satisfaisant au délai un. Dans le tableau 5.4, le test de El Himdi et Roy (1997) sur les résidus standardisés suggère l'existence de la causalité en moyenne aux délais 0 et -1. Avant de poursuivre avec le test de causalité, nous augmenterons les modèles afin d'éliminer toute causalité en moyenne. Le modèle augmenté suivant à été retenu (voir les coefficients estimés dans le tableau 5.3) :

$$
\begin{aligned}
\mathbf{X}_{1t} &= \boldsymbol{\epsilon}_{1t}, \\
(\mathbf{I}_2 - \boldsymbol{\Phi}^{(2)}B)\mathbf{X}_{2t} &= \boldsymbol{\Lambda}_0^{(2)}\mathbf{X}_{1t} + \boldsymbol{\Lambda}_1^{(2)}\mathbf{X}_{1,t-1} + (\mathbf{I}_2 + \boldsymbol{\Theta}^{(2)}B + \boldsymbol{\Theta}^{(3)}B^2)\boldsymbol{\epsilon}_{2t}.
\end{aligned}
$$

Toujours dans le tableau 5.4 on peut remarquer que l'augmentation des modèles avec des variables exogènes dans le modèle a éliminé complètement la causalité en moyenne. Nous pouvons effectuer le test de causalité en variance. Les tableaux 5.5 et 5.6 présentent les résultats de nos tests de causalité en variance reposant sur les tests $\mathbf{ER}^*_{nM}$ et $\mathbf{LL}^*_{nM}$, $M = 1, \ldots, 12$, et les tests $\mathbf{ER}^*_n(k)$ et $\mathbf{LL}^*_n(k)$, $k = -10, \ldots, 10$, respectivement.

Les deux tests sont significatifs (aussi bien à 1% qu'à 5%) pour les délais globaux. Il semble alors exister une relation de causalité en variance entre ces deux séries d'indices. Les tests de délais individuels nous renseignent un peu plus sur la direction de la causalité. Les deux tests s'accordent sur la causalité instantanée (délai 0) ainsi qu'au délai -1. On peut alors conclure que les indices européens causent les indices asiatiques. Les deux tests sont aussi significatifs à 5% pour le délai positif $k = 2$. On serait donc en situation de causalité dans les deux sens, les indices asiatiques causeraient également les indices européens. Notons tout de même cette discordance des deux tests aux délais 9 et 10 ; le test $\mathbf{ER}^*_n(k)$ est significatif à 5% contrairement au test $\mathbf{LL}^*_n(k)$. Ce pourrait possiblement s'expliquer en partie par la perte de puissance du test selon l'approche de Ling et Li (1997) comme nous le signalions déjà dans l'étude menée dans l'article.

Tableau 5.1 – Estimateurs pour les modèles des indices asiatiques et européens avec les valeurs t correspondantes.

	Indices asiatiques			Indices Européens		
	Estimateurs	Valeurs-t	Valeurs-P	Estimateurs	Valeurs-t	Valeurs-P
$\hat{C}_{11}$	1.172×10^{-5}	3.824	0.0001	8.349×10^{-6}	4.231	0^+
$\hat{C}_{21}$	9.927×10^{-6}	3.370	0.0007	7.853×10^{-6}	4.045	0^+
$\hat{C}_{22}$	1.441×10^{-6}	3.551	0.0004	8.550×10^{-6}	4.028	0^+
$\hat{A}_{11}$	0.129	6.653	0^+	0.122	7.994	0^+
$\hat{A}_{21}$	0.085	6.362	0^+	0.105	7.512	0^+
$\hat{A}_{22}$	0.108	6.971	0^+	0.105	6.852	0^+
$\hat{B}_{11}$	0.833	35.772	0^+	0.834	40.586	0^+
$\hat{B}_{21}$	0.867	39.130	0^+	0.852	41.809	0^+
$\hat{B}_{22}$	0.860	44.560	0^+	0.862	43.005	0^+

Tableau 5.2 – Tests statistiques aux délais individuels appliqués sur les résidus standardisés.

	$\mathrm{H}_n^{(ii)}(k)$		$\mathrm{HM}_n^{(ii)}(k)$	
k	$\mathbf{X}_{1t}$	$\mathbf{X}_{2t}$	$\mathbf{X}_{1t}$	$\mathbf{X}_{2t}$
1	15.2225	20.1658	15.3071	21.4544
2	8.4826	5.8189	8.8968	6.2492
3	6.1840	3.0242	6.2415	3.1735
4	5.6438	4.0888	5.7151	4.3435
5	8.8470	1.8379	8.9797	1.9512
6	8.0766	1.1667	8.3517	1.2675
7	2.3754	3.4558	2.4533	3.5679
8	4.2873	4.2037	4.4314	4.4498
9	2.0760	4.4327	2.1804	4.6192
10	8.5685	9.7098	9.0504	10.6733
11	7.1816	2.8381	7.6422	3.0896
12	2.6542	6.0852	2.7432	6.4172

Tableau 5.3 – Estimateurs des paramètres pour le modèle augmenté de l'indice asiatique avec les valeurs t correspondantes.

	Estimateurs	Valeurs-t	Valeurs-P		Estimateurs	Valeurs-t	Valeurs-P
$\hat{\Phi}_{11}^{(2)}$	-0.2169	-0.90983	0.3632	$\hat{C}_{1,11}^{(2)}$	8.710×10^{-6}	3.36646	0.0008
$\hat{\Phi}_{21}^{(2)}$	0.6921	3.12052	0.0018	$\hat{C}_{1,21}^{(2)}$	5.781×10^{-6}	3.09188	0.0020
$\hat{\Phi}_{12}^{(2)}$	0.2367	0.89203	0.3727	$\hat{C}_{1,22}^{(2)}$	7.663×10^{-6}	3.00219	0.0027
$\hat{\Phi}_{22}^{(2)}$	-0.7396	-3.07852	0.0021	$\hat{A}_{1,11}^{(2)}$	0.1428	5.81730	0^+
$\hat{\Theta}_{11}^{(2)}$	-0.07949	-0.32105	0.7483	$\hat{A}_{1,21}^{(2)}$	0.06748	4.87288	0^+
$\hat{\Theta}_{21}^{(2)}$	-0.8922	-3.81130	0.0001	$\hat{A}_{1,22}^{(2)}$	0.08793	6.77659	0^+
$\hat{\Theta}_{12}^{(2)}$	-0.2219	-0.82578	0.4092	$\hat{B}_{1,11}^{(2)}$	0.8143	28.68972	0^+
$\hat{\Theta}_{22}^{(2)}$	0.5750	2.36098	0.0185	$\hat{B}_{1,21}^{(2)}$	0.8773	37.98448	0^+
$\hat{\Theta}_{11}^{(3)}$	0.07492	1.48452	0.1381	$\hat{B}_{1,22}^{(2)}$	0.8879	51.71745	0^+
$\hat{\Theta}_{21}^{(3)}$	0.1151	1.82640	0.0682				
$\hat{\Theta}_{12}^{(3)}$	0.02442	0.44716	0.6549				
$\hat{\Theta}_{22}^{(3)}$	-0.1522	-2.41256	0.0161				
$\hat{\Lambda}_{0,11}^{(2)}$	0.2801	4.90100	0^+				
$\hat{\Lambda}_{0,21}^{(2)}$	0.2281	3.05020	0.0023				
$\hat{\Lambda}_{0,12}^{(2)}$	0.1903	3.90918	0.0001				
$\hat{\Lambda}_{0,22}^{(2)}$	0.3635	5.60075	0^+				
$\hat{\Lambda}_{1,11}^{(2)}$	0.06559	0.94773	0.3436				
$\hat{\Lambda}_{1,21}^{(2)}$	0.005542	0.06113	0.9513				
$\hat{\Lambda}_{1,12}^{(2)}$	0.3868	5.29883	0^+				
$\hat{\Lambda}_{1,22}^{(2)}$	0.5329	6.50533	0^+				

Tableau 5.4 – Test statistique de El Himdi et Roy (1997) pour tester la causalité en moyenne. Les valeurs P approximatives sont également données.

	Modèles originaux		Modèles augmentés	
k	$\mathrm{HM}_n^{(12)}(k)$	valeurs-P	$\mathrm{HM}_n^{(12)}(k)$	valeurs-P
-10	6.1113	0.1910	3.3504	0.5010
-9	7.7648	0.1006	7.7714	0.1003
-8	2.9323	0.5692	4.6144	0.3292
-7	5.9058	0.2063	4.7062	0.3188
-6	5.8284	0.2123	6.2013	0.1846
-5	8.1235	0.0872	6.8110	0.1462
-4	1.7019	0.7904	2.7418	0.6019
-3	2.1988	0.6993	0.8272	0.9348
-2	0.5790	0.9654	0.5003	0.9735
-1	121.8829	0.0000	1.6776	0.7948
0	155.4515	0.0000	2.3557	0.6707
1	1.2145	0.8757	4.1869	0.3813
2	0.8327	0.9340	2.5280	0.6396
3	1.6972	0.7912	3.3506	0.5010
4	7.9082	0.0950	7.8428	0.0975
5	3.6483	0.4557	1.5731	0.8136
6	2.9422	0.5676	0.4118	0.9815
7	3.0425	0.5507	0.7028	0.9510
8	6.5674	0.1606	4.6992	0.3196
9	3.1816	0.5279	5.0808	0.2791
10	8.2520	0.0828	3.3157	0.5064

Tableau 5.5 – Tests statistiques globaux pour tester la causalité en variance $\mathrm{ER}^*_{n,M}$ et $\mathrm{LL}^*_{n,M}$, $M = 1, \ldots, 10$. Les degrés de liberté et les valeurs P sont également fourni.

	Test statistique $\mathrm{ER}^*_{n,M}$			Test statistique $\mathrm{LL}^*_{n,M}$		
M	$\mathrm{ER}^*_{n,M}$	$d.f.$	valeurs-P	$\mathrm{LL}^*_{n,M}$	$d.f.$	valeurs-P
1	85.5434	27	0^+	35.3103	3	0^+
2	118.5958	45	0^+	44.6479	5	0^+
3	150.0553	63	0^+	47.9847	7	0^+
4	161.8508	81	0^+	48.4233	9	0^+
5	194.3549	99	0^+	54.8674	11	0^+
6	204.4938	117	0^+	57.9993	13	0^+
7	212.0027	135	0^+	58.9991	15	0^+
8	228.0824	153	0.0001	62.3724	17	0^+
9	259.9222	171	0^+	65.2422	19	0^+
10	294.3717	189	0^+	67.6888	21	0^+

Tableau 5.6 – Tests statistiques aux délais individuels pour détecter la causalité en variance $\mathrm{ER}_n^*(k)$ et $\mathrm{LL}_n^*(k)$, $|k| \leq 10$.

	Test statistique $\mathrm{ER}_n^*(k)$		Test statistique $\mathrm{LL}_n^*(k)$	
k	$\mathrm{ER}_n^*(k)$	valeurs-P	$\mathrm{LL}n^*(k)$	valeurs-P
-10	4.4454	0.8797	0.0497	0.8237
-9	7.3169	0.6042	1.7364	0.1876
-8	10.3740	0.3211	3.0840	0.0791
-7	1.9074	0.9928	0.0003	0.9853
-6	4.3576	0.8863	0.1295	0.7189
-5	15.4243	0.0799	2.7595	0.0967
-4	8.7248	0.4631	0.0292	0.8644
-3	14.9834	0.0914	1.7625	0.1843
-2	15.5564	0.0767	3.9732	0.0462
-1	46.2570	0^+	13.2372	0.0003
0	28.6681	0.0007	22.0715	0^+
1	10.6184	0.3028	0.0016	0.9681
2	17.4959	0.0415	5.3644	0.0206
3	16.4762	0.0576	1.5743	0.2096
4	3.0707	0.9614	0.4094	0.5223
5	17.0798	0.0475	3.6847	0.0549
6	5.7813	0.7616	3.0023	0.0831
7	5.6015	0.7790	0.9994	0.3174
8	5.7057	0.7690	0.2894	0.5906
9	24.5229	0.0035	1.1334	0.2870
10	30.0041	0.0004	2.3969	0.1216

Conclusion

Dans ce mémoire, nous nous sommes intéressé aux tests de causalité en variance entre deux séries chronologiques dans un contexte multivarié. Dans un premier temps, nous avons généralisé avec une première approche le test de causalité en variance développé par Cheung et Ng (1996). En nous appuyant sur le travail de El Himdi et Roy (1997) et Duchesne (2004), nous avons proposé une statistique de test basée sur la matrice des corrélations croisées des résidus standardisés et les produits croisés des résidus standardisés. Dans une deuxième approche, utilisant la transformation proposée dans Ling et Li (1997), nous avons construit un test statistique à partir des résidus transformés. Ce deuxième test est attrayant en pratique du fait de sa simplicité. En effet sa mise en oeuvre est aisée avec les logiciels statistiques courants. Cependant cette simplicité a un coût ; dans la mesure où cette transformation peut entrainer une perte en puissance. Nous avons pour chacune des deux approches proposé des tests de type portemanteau et des tests de délais individuels, en spécifiant à chaque fois la distribution asymptotique de la statistique de test sous l'hypothèse nulle d'absence de causalité en variance. Les tests aux délais individuels permettent le cas échéant de caractériser la direction de la causalité. Des versions modifiées des tests précédents ont été proposé dans l'optique d'améliorer les propriétés empiriques.

Dans un deuxième temps, nous avons effectué des simulations Monte Carlo. Une étude de niveau a montré que le niveau empirique de chacun des quatre tests est raisonnablement proche du niveau théorique. Une étude de puissance quant à elle nous a révélée que le test issu de la deuxième approche, malgré la simplicité de sa mise en oeuvre est en général moins puissant que le premier test, au moins pour les modèles considérés dans ce travail.

Dans un troisième temps, afin d'illustrer nos méthodes, nous avons appliqué ces tests à des données réelles via deux études de cas. Après avoir fait le test de causalité en moyenne en utilisant la méthode de El Himdi et Roy, nous avons effectué le test de causalité en variance entre deux indices boursiers nord-américains et deux indices européens d'une part, et d'autre part entre deux indices européens et deux indices asiatiques.

Bibliographie

Ansley, C. F. et Newbold, P. (1979), 'On the finite sample distribution of residual autocorrelations in autoregressive-moving average models', *Biometrika*, **66**, 547–553.

Berndt, E., Hall, B., Hall, R. et Hausman, J. (1974), 'Estimation and inference in nonlinear structural models', *Annals of Social Measurement* **3**, 653–665.

Bollerslev, T. P. (1986), 'Generalised autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**, 309–28.

Bollerslev, T. P., Engle, R. F. et Nelson, D. B. (1994), 'ARCH models', dans *Handbook of Econometrics*, Volume 4 (Engle, R. F. et Mcfadden, D. L., Éditeurs), Amsterdam : Elsevier Science, 2959–3038.

Bollerslev, T. P., Engle, R. F. et Wooldridge, J. M. (1988), 'A capital asset pricing model with time varying covariances', *Journal of Political Economy* **96**, 116–31.

Brockwell, P. J. et Davis, R. A. (1991), *Time Series : Theory and Methods*, Deuxième Édition, Springer-Verlag : New York.

Caporin, M. (2007), 'Variance (non) causality in multivariate GARCH', *Econometric Reviews* **26**, 1–24.

Cheung, Y.-W. et Ng, L. K. (1990), 'The dynamics of S&P 500 index and S&P 500 futures intraday price volatilities', *Review of Futures Markets* **9**, 458–486.

Cheung, Y.-W. et Ng, L. K. (1996), 'A causality-in-variance test and its application to financial market prices', *Journal of Econometrics* **72**, 33–48.

Comte, F. et Lieberman, O. (2000), 'Second-order noncausality in multivariate GARCH processes', *Journal of Time Series Analysis* **21**, 535–557.

Comte, F. et Lieberman, O. (2003), 'Asymptotic theory for multivariate GARCH processes', *Journal of Multivariate Analysis* **84**, 61–84.

Ding, Z. (1994), 'Time series analysis of speculative returns', dissertation, University of California at San Diego.

Duchesne, P. (2004), 'On matricial measures of dependence in vector ARCH models with applications to diagnostic checking', *Statistics & Probability Letters* **68**, 149–160.

Duchesne, P. et Lalancette, S. (2002), 'A test for multivariate autoregressive conditional heteroskedasticity effects in vector time series models', *Les Cahiers du GERAD* G-2002-45, GERAD, août 2002.

Duchesne, P. et Lalancette, S. (2003), 'On testing for multivariate ARCH effects in vector time series models', *The Canadian Journal of Statistics* **31**, 275–292.

El Himdi, K. et Roy, R. (1997), 'Tests for noncorrelation of two multivariate ARMA time series', *The Canadian Journal of Statistics* **25**, 233–256.

Engle, R. F. (1982), 'Autoregressive conditional heteroskedasticity with estimates of the variance of U. K. inflation', *Econometrica* **50**, 987–1007.

Engle, R. F. et Kroner, K. F. (1995), 'Multivariate simultaneous generalized ARCH', *Econometric Theory* **11**, 122–50.

Francq, C. et Zakoian, J.-M. (2009), *Modèles GARCH*, Economica : Paris.

Francq, C. et Zakoian, J.-M. (2010), 'QML estimation of a class of multivariate GARCH models without moment conditions on the observed process', document de travail, MPRA Papier No. 20779, février 2010.

Fuller, W. A. (1996), *Introduction to Statistical Time Series*, Deuxième Édition, New York : Wiley.

Granger, C. W. J., Robins, R. P. et Engle, R. F. (1986), 'Wholesale and retail prices : Bivariate time-series modeling with forecastable error variances', dans Belsley, D. A. and Kuh, E., Editeurs, *Model Reliability*, MIT Press, Cambridge, MA, 1–17.

Granger, C. W. J. (1969), 'Investigating causal relations by econometric methods and cross-spectral methods', *Econometrica* **34**, 424–438.

Hafner, C. M. et Herwartz, H. (2006), 'A Lagrange multiplier test for causality in variance', *Economics Letters* **93**, 137–141.

Hafner, C. M. et Herwartz, H. (2008), 'Testing for causality in variance using multivariate GARCH models', *Annales d'Économie et de Statistique* **89**, 215–241.

Hafner, C. M. et Preminger, A. (2009), 'On asymptotic theory for multivariate GARCH models', *Journal of Multivariate Analysis* **100**, 2044–2054.

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press : Princeton, New Jersey.

Harville, D. A. (1997), *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag : New York.

Haugh, L. D. (1976), 'Checking the independence of two covariance-stationary time series : A univariate residual cross-correlation approach', *Journal of the American Statistical Association* **71**, 378–385.

Hiemstra, C. et Jones, J. D. (1994), 'Testing for linear and nonlinear Granger causality in the stock price-volume relation', *Journal of Finance* **49**, 1639–1664.

Hong, Y. (2001), 'A test for volatility spillover with application to exchange rates', *Journal of Econometrics* **103**, 183–224.

Hosking, J. R. M. (1980), 'The multivariate portmanteau statistic', *Journal of the American Statistical Association* **75**, 602–608.

Jacobs R. L., Leamer E. E. et Ward M. P. (1979), 'Difficulties in testing for causation', *Economic Inquiry* **17**, 401–413.

Li, W. K. (2004), *Diagnostic Checks in Time Series*, Chapman & Hall/CRC : New York.

Li, W. K. et Mak, T. K. (1994), 'On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity', *Journal of Time Series Analysis* **15**, 627–636.

Ling, S. et Li, W. K. (1997), 'Diagnostic checking of nonlinear multivariate time series with multivariate ARCH errors', *Journal of Time Series Analysis* **18**, 447–464.

Ling, S. et McAleer, M. (2003), 'Asymptotic theory for a vector ARMA-GARCH model', *Econometric Theory* **19**, 280–310.

Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag : New York.

McLeod, A. I. et Li, W. K. (1983), 'Diagnostic checking ARMA time series models using squared-residual autocorrelations', *Journal of Time Series Analysis* **4**, 269–273.

Poulin, J. et Duchesne, P. (2008), 'On the power transformation of kernel-based tests for serial correlation in vector time series : some finite sample results and a comparison with the bootstrap', *Computational Statistics and Data Analysis* **52**, 4432–4457.

Ross, S. A. (1989), 'Information and volatility : The no-arbitrage martingale approach to timing and resolution irrelevancy', *Journal of Finance* **44**, 1–17.

Sims, C. (1972), 'Money, income and causality', *American Economic Review* **62**, 540- 552.

Stout, W. (1974), *Almost Sure Convergence*, New York : Academic Press.

Tse, Y. K. (2002), 'Residual-based diagnostics for conditional heteroscedasticity models', *Econometrics Journal* **5**, 358–373.

Tse, Y. K. et Tsui, A. K. C. (1999), 'A note on diagnosing multivariate conditional heteroscedasticity models', *Journal of Time Series Analysis* **20**, 679–691.

Wong, H. et Li, W. K. (1996), 'Distribution of the cross-correlations of squared residuals in ARIMA models', *The Canadian Journal of Statistics* **24**, 489–502.

Wong, H. et Li, W. K. (2002), 'Detecting and diagnostic checking multivariate conditional heteroscedastic time series models', *Annals of the Institute of Statistical Mathematics* **54**, 45–59.

Zivot, E. et Wang, J. (2003), *Modeling Financial Time Series with S-PLUS*, Springer-Verlag : New York.