

Université de Montréal

**La prévisibilité des rendements financiers : une
analyse de bootstrap**

Par
Hasina Rasata

Directrice de recherche : Silvia Gonçalves

Centre de Documentation
Dép. de sciences économiques
Université de Montréal
C. P. 8128, Succ. "A"
Montréal, Q. B., Canada, H3C 2J4

**Département de Sciences économiques
Faculté des Arts et des Sciences**

**Rapport de recherche présenté à la Faculté des Études Supérieures en vue de
l'obtention du grade de Maîtrise ès sciences (M. Sc)**

Septembre, 2002

REMERCIEMENTS

J'exprime ma haute considération à Madame Silvia Gonçalves pour l'intérêt qu'elle a porté à ce rapport de recherche, et pour ses bienveillants conseils et critiques. Je lui adresse ma profonde gratitude.

Je présente aussi mes meilleurs remerciements à tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

Je remercie Dieu tout puissant pour la force qu'Il m'a prêté dans la réalisation de ce travail.

Je dédie ce fruit de mon travail à ma mère bien-aimée.

RÉSUMÉ

La prévisibilité des rendements financiers constitue un domaine de recherche intéressant en économie financière. Les tests y afférent rencontrent au moins un problème statistique : les données financières sont hétéroscédastiques. La plupart des travaux réalisés dans ce sens ont tendance soit à ignorer le fait empirique que les données financières sont hétéroscédastiques, soit à y imposer une forme particulière. Ces faits pourraient affecter les résultats des tests effectués. Dans ce rapport, nous utilisons une méthode robuste à l'hétéroscédasticité conditionnelle : le wild bootstrap récursif. Le wild bootstrap récursif, comme toute approche de bootstrap, nous permet d'avoir des résultats plus précis que la théorie asymptotique pour des petits échantillons. La statistique de test utilisée est celle du ratio de variance. Nous considérons deux approches, l'une est univariée tandis que l'autre est multivariée. Les résultats empiriques pour les deux approches n'établissent aucune évidence sur la prévisibilité des rendements financiers.

SUMMARY

The predictability of stock returns constitutes an interesting field of research in financial economics. Major tests performed to assess that predictability suffer at least one statistical problem: that is, financial data are heteroskedastic and researchers often modelize it in a particular way. This could affect those tests' results. This research paper aims at providing a method robust to conditional heteroskedasticity: the recursive wild bootstrap. The recursive wild bootstrap, like any bootstrap approach, provides more precise results than the asymptotic theory for small samples. The test statistic performed is the variance ratio statistic. We consider two approaches: one univariate and the other multivariate. Our empirical results from both approaches establish no evidence on the predictability of stock returns.

TABLE DES MATIÈRES

REMERCIEMENTS	i
RÉsumÉ	ii
Summary	iii
Liste des tableaux	v
Liste des figures	vi
Introduction	1
1. La prévisibilité des rendements d'actifs financiers	4
1.1. Importances	4
1.2. Relations algébriques	5
1.3. Études empiriques	8
2. Le bootstrap	13
2.1. Un exemple de bootstrap: le cas de la moyenne empirique	13
2.2. Le test bootstrap classique	16
2.3. Le wild bootstrap (WB) récursif	18
3. Le test de prévisibilité des rendements financiers par le WB récursif	21
3.1. Le modèle	21
3.2. Le ratio de variance (variance ratio statistics)	23
3.2.1. Le ratio de variance univarié	24
3.2.2. Le ratio de variance multivarié	26
3.3. La démarche du test	28
4. Analyse empirique	31
4.1. Description des données	31
4.2. Résultats	32
Conclusion	43
BIBLIOGRAPHIE	45
ANNEXE I	47
ANNEXE II	48

LISTE DES TABLEAUX

Tableau 1 : Statistiques descriptives des variables.....	32
Tableaux 2 : « Equal-tailed 95% percentile bootstrap confidence interval » pour $M_r(k)$	33
Tableaux 3 : Estimation du modèle VAR de log rendements, de ratio dividende prix et de rendement relatif de bons de trésor.....	37
Tableaux 4 : « Equal-tailed 95% percentile bootstrap confidence interval » pour $VR(k)$ Échantillon : février 1927 à novembre 1987	41

LISTE DES FIGURES

Figure 1: Ratios de variance de log rendements pour l'horizon k , $M_r(k)$	35
Figure 2: Ratios de variance pour l'horizon k dérivés à partir du modèle VAR, $VR(k)$	40

INTRODUCTION

La prévisibilité des rendements d'actifs financiers est au coeur de la théorie financière moderne. Personne ne doute de son importance dans l'allocation des portefeuilles. Notons, par exemple, que les modèles d'évaluation d'actifs tels le CAPM (Capital Asset Pricing Model) et le CCPAM (Consumption-based Capital Asset Pricing Model) sont basés sur l'idée de variabilité des rendements espérés des actifs financiers vus leurs risques respectifs. Dans ce présent rapport, nous essayerons de répondre à la question de prévisibilité des rendements d'actifs financiers en utilisant une méthode à la fois robuste à l'hétéroscédasticité conditionnelle et sans forme particulière pour l'hétéroscédasticité. La façon pour faire l'inférence est ici une manière de tester la forme faible de l'hypothèse de l'efficience des marchés des actifs.

Depuis quelques décennies, une nouvelle vague de pensées révolutionne la théorie des marchés efficients. Celle-ci adhère l'idée selon laquelle la prévisibilité des rendements d'actifs financiers n'est pas forcément incompatible avec la théorie des marchés efficients. La variabilité dans les taux de rendements espérés des actifs financiers est nécessaire pour rémunérer les risques liés à ces actifs.

Les économistes ont essayé de dériver des relations algébriques liant les rendements espérés à des variables macroéconomiques et/ou financières. Campbell et Shiller (1988) ont montré l'existence d'une relation qui lie le logarithme du ratio dividende-prix aux rendements excédentaires espérés. Une telle découverte est devenue, par la suite, une justification du souci de tester empiriquement, la puissance du ratio dividende-prix à expliquer voire prédire les rendements financiers.

Deux questions se posent ici :

- 1) Comment traiterons-nous, dans ce cas, les actifs qui ne paient pas de dividendes?
- 2) Comment considérons-nous le fait que les données financières cachent de l'hétéroscédasticité conditionnelle? Les modèles de type ARCH (AutoRegressive Conditional Heteroscedasticity), qui émergent avec Engle au début des années 80, ont été développés pour tenir compte de cette hétéroscédasticité conditionnelle dans les données financières.

Une manière de répondre à la première question est de dire que vue la modernisation de la gestion des entreprises actuellement, la plupart des actifs paient de dividendes soit de façon explicite (en liquide), soit de façon déguisée (ex : en actions de la société). Tandis que pour la deuxième question, il est à signaler que l'un des objectifs de ce présent rapport est de considérer l'existence de l'hétéroscédasticité conditionnelle dans les données financières, au moment où l'on fait l'inférence, sans avoir à imposer une forme particulière. Pour ce faire, nous utiliserons une méthode de bootstrap robuste à la présence d'hétéroscédasticité, la méthode « wild bootstrap récursif ».

Une grande partie de la littérature, couvrant ce sujet particulier, ne considèrent pas l'hétéroscédasticité (ex : Goetzmann et Jorion, 1993). D'autres la traitent de façon paramétrique (ex : Hodrick, 1992).

Une façon de tester la prévisibilité des rendements à partir des rendements passés est d'utiliser le test de ratio de variance (variance ratio test). Ce test utilise le simple fait que la variance de la somme des variables aléatoires non corrélées est égale à la somme des variances individuelles. En utilisant de tel test à la fois univarié et multivarié, nous ne pouvons pas rejeter l'hypothèse nulle d'absence d'autocorrélation

dans les rendements financiers. Autrement dit, les rendements d'actifs financiers sont, stastiquement, non prévisibles à l'aide des rendements passés. Notons que la prise de décision est faite en construisant des intervalles de confiance bootstrap, où chaque intervalle est constitué à partir de 1000 simulations, pour la statistique ratio de variance et en regardant si ces intervalles contiennent l'unité ou non. Si la valeur un est à l'intérieur de l'intervalle de confiance alors les rendements sont imprévisibles.

Le plan de notre travail comporte quatre sections. À la première section, nous présenterons les motivations de notre rapport ainsi qu'une brève revue de la littérature. La section deux sera consacrée à quelques rappels utiles sur la procédure de bootstrap. La section trois, quant à elle, décrira les démarches des tests effectués. Nous présenterons à la dernière section nos résultats empiriques et nos commentaires.

1. La prévisibilité des rendements d'actifs financiers

La prévisibilité des rendements d'actifs financiers est devenue le leitmotiv de la théorie financière moderne. Plusieurs études ont été faites pour tester s'il y a ou non de prévisibilité des rendements. Une question se pose ici, pourquoi la prévisibilité est-elle importante du point de vue économique?

1.1. Importances

Depuis plusieurs décennies, différents auteurs ont rejeté l'idée d'incompatibilité entre la théorie des marchés efficients¹ et la prévisibilité des rendements. En effet, la prévisibilité des rendements financiers peut être expliquée par la variation à travers le temps dans le taux auquel les investisseurs rationaux escomptent ses revenus futurs à partir des actifs risqués. Le modèle CCPAM (Consumption-based Capital Asset Pricing Model)² implique, par exemple, que les prix des actifs ne se comportent comme des martingales qu'en cas de neutralité au risque : une situation généralement irréaliste. Dans ce cadre, la variabilité dans les rendements espérés futurs rémunère

¹ Un marché est efficient lorsque le prix s'y forme sur la base de l'ensemble de l'information disponible. On peut distinguer trois types de marché efficient qui sont fonction du type d'information disponible :

- L'hypothèse de marché efficient en forme faible qui explicite que les prix reflètent toute l'information contenue dans la série historique des prix.
- L'hypothèse de marché efficient en forme semi forte établit que les prix reflètent toute l'information publique disponible.
- L'hypothèse de marché efficient en forme forte qui établit que toute l'information connue, publique et privée, est reflétée dans les prix du marché.

² Le modèle CCPAM est présenté comme suit :

$$P = \delta E_t \left[\frac{U_1(C_{t+1})}{U_1(C_t)} (P_{t+1} + Y_{t+1}) \right] \Leftrightarrow 1 = \delta E_t \left[\frac{U_1(C_{t+1})}{U_1(C_t)} (1 + R_{t+1}) \right]$$

où δ est le taux d'escompte, P_t est le prix de l'actif en t et $U_1(C_t)$ est l'utilité marginale de la consommation en t et R_{t+1} est le rendement de l'actif en $t+1$.

les risques liés aux actifs. Ainsi, la prévisibilité des rendements n'est qu'un outil pour comprendre le mécanisme du marché.

Notons aussi que la prévisibilité des rendements financiers permet aux agents de changer l'allocation de leur portefeuille entre les actifs de façon plus performante que de prendre les rendements espérés inconditionnels. Prenons à titre d'exemple, le cas du modèle CAPM (Capital Asset Pricing Model)³, sous ses hypothèses restrictives, la proportion d'un actif dans un portefeuille risqué devrait être proportionnel aux rendements espérés excédentaires du marché. Donc, lorsque les rendements espérés varient dans le temps, on voudrait augmenter le poids des actifs dont les rendements sont prévus être élevés dans le futur (et vice versa) en supposant que les variances sont constantes.

Bref, la variabilité des rendements espérés joue un grand rôle dans la mise en place d'une stratégie optimale d'investissements. Conséquemment, plusieurs études ont été effectuées pour déduire les soubassements théoriques de la prévisibilité des rendements d'actifs financiers.

1.2. Relations algébriques

Une grande partie de la nouvelle littérature sur l'économie financière s'efforce à dériver une (des) relation(s) entre les rendements espérés futurs et les variables macroéconomiques et/ou financières. Celle-ci conforte l'idée de la possibilité de prévoir les rendements des actifs financiers.

³ Le modèle CAPM peut être présenté par : $E(R_i) = R_f + \beta_{im} [E(R_m) - R_f]$

où R_m est le rendement du portefeuille de marché, R_f est le rendement sans risque, R_i est le rendement de l'actif risqué i et $\beta_{im} = \frac{\text{cov}(R_i, R_m)}{\text{var}(R_m)}$

Lettau et Ludvigson (2001) ont montré, en l'occurrence, que l'utilisation d'un proxy du logarithme du ratio consommation richesse agrégée, est adéquate pour prévoir les rendements.

Notons que, la richesse agrégée, W_t , est définie comme la somme du capital humain, H_t et de la richesse financière, A_t . Lettau et Ludvigson (2001a) ont argumenté que la contrainte budgétaire standard implique que le logarithme de la consommation, c_t , le logarithme du revenu de travail, y_t et le logarithme de la richesse financière, a_t sont cointégrés. Ainsi, les déviations par rapport à la tendance commune dans ces variables peuvent être imputables à des fluctuations dans le logarithme du ratio consommation richesse agrégée, cay_t , qui est une variable prédictrice des rendements financiers. Par ailleurs, le capital humain, H_t est inobservable. Lettau et Ludvigson assument que le revenu du travail agrégé, qui est non stationnaire, définit la tendance dans le capital humain. D'où l'écriture du logarithme du capital humain, $h_t = y_t + z_t$, où z_t est une variable aléatoire stationnaire.

En utilisant cette relation, ils ont dérivé l'équation du cay_t telle que :

$$cay_t \equiv c_t - \alpha_a a_t - \alpha_y y_t \approx k + E_t \left[\sum_{i=1}^{\infty} \rho_w^i (r_{w,t+i} - \Delta c_{t+i}) + (1-w) z_t \right] \quad (1)$$

où

r_w est le log du rendement de la richesse agrégée,

ρ_w est le ratio à l'état stationnaire des nouveaux investissements et de la richesse totale :

$$\rho_w = \frac{(W - C)}{C} \quad (2)$$

et k est une constante.

L'équation (1) est obtenue en appliquant un développement d'ordre 1 de Taylor à l'équation dynamique de la richesse agrégée :

$$W_{t+1} = (1 + R_{w,t+1})(W_t - C_t) \quad (3)$$

L'équation (1) nous dit qu'une variation du ratio consommation-richesse agrégée est imputable soit aux rendements, soit à l'accroissement de la consommation, soit à l'effet combiné des deux. Ainsi, les agents auront une tendance à augmenter la valeur de leur portefeuille quand le ratio consommation-richesse agrégée est élevé. Il est à noter ici que, dans la plupart des cas (Fama et French, 1988; Campbell et Shiller, 1988; Hodrick, 1992; Goetzmann et Jorion, 1993; Lo et Mackinlay, 1997), les auteurs s'intéressent à tester la puissance du ratio dividende-prix à prédire les rendements d'actifs financiers.

Campbell et Shiller (1988) ont montré qu'une approximation pour l'expression du logarithme du ratio dividende-prix peut-être écrite comme suit :

$$p_t - d_t \approx k + E_t \sum_{j=1}^{\infty} \rho_s^j (\Delta d_{t+j} - r_{s,t+j}) \quad (4)$$

où ρ_s est égal à l'inverse du ratio dividende-prix, quand celui-ci est constant, augmenté par l'unité.

$$\rho_s = \frac{P}{(P + D)} \quad (5)$$

et k est une constante qui ne joue aucun rôle dans l'analyse. Cette équation représente le « modèle dynamique de croissance de dividende » et qui est obtenu par une

transformation de Taylor de premier ordre de l'équation définissant le log de rendement:

$$r_{st} = \log(P_t + D_t) - \log(P_t) \quad (6)$$

En prenant l'espérance, l'équation ci-dessus nous dit que lorsque le ratio dividende-prix est élevé, les agents vont anticiper des rendements plus élevés de ses portefeuilles dans le futur ou un faible taux de croissance des dividendes. Campbell (1991) et Cochrane (1991a) ont conclu que toute variation dans le log de ratio dividende-prix, $p_t - d_t$, est attribuable non à la variation dans la croissance des dividendes espérés futurs mais aux changements des prévisions des rendements excédentaires. Bref, l'existence des relations algébriques liant les rendements espérés et les différentes variables macroéconomiques et/ou financières, particulièrement le ratio dividende-prix, justifie le souci de tester empiriquement la puissance de ce dernier à expliquer voire prédire les rendements financiers.

1.3. Études empiriques

Depuis Fama et French (1988), beaucoup d'études ont essayé de tester le pouvoir prédictif des ratios dividende-prix sur les rendements financiers. Le modèle de base le plus utilisé est le modèle de régression linéaire tel que:

$$\log(R_{t+k,k}) = \alpha_k + \beta_k \left(\frac{D_t}{P_t} \right) + \varepsilon_{t+k,k} \quad (7)$$

où l'hypothèse nulle de non prévisibilité revient à tester si $\beta_k=0$ dont k représente l'horizon de la prévision.

Ici,

$$\log(R_{t+k,k}) = \log(R_{t+1}) + \dots + \log(R_{t+k}) \quad (8)$$

est le rendement continûment composé de k périodes dont

$$R_{t+1} = \frac{P_{t+1} + d_{t+1}}{P_t} \quad (9)$$

avec P_t est le prix réel à la fin du mois t de l'actif et d_t est le dividende réel payé durant le mois t .

D_t représente la série des dividendes annualisés. Cette dernière est utilisée afin de tenir compte de la présence de saisonnalité dans la série des dividendes.

$$D_t = d_t + (1+r_t)d_{t-1} + \dots + (1+r_t)(1+r_{t-1})\dots(1+r_{t-10})d_{t-11} \quad (10)$$

avec r_t est le rendement du bon de trésor d'un mois.

Une façon de faire l'inférence sur β_k , dans le cas de présence de dépendance et d'hétéroscédasticité des observations, est d'estimer la distribution échantillonnale de $\hat{\beta}_k$ sous la nulle et d'utiliser cette distribution estimée pour attacher une p-value à la valeur observée de $\hat{\beta}_k$ ⁴.

Pour ce faire, on recourt à un processus générateur de données qui impose la nulle à savoir le modèle VAR (Vector AutoRegressive model).

⁴ Autrement dit, une façon de faire l'inférence sur β_k est d'utiliser une approche de bootstrap.

L'utilisation du modèle VAR(1), est initiée par Campbell et Shiller (1988b), avec au moins deux variables: le logarithme des rendements et le ratio dividende-prix (dividend yields). Notons que dans les différentes études sur le pouvoir prédictif du ratio dividende-prix utilisant le modèle VAR(1), les auteurs soit ils ne considèrent pas la présence d'hétéroscédasticité dans les données financières, soit ils imposent une forme particulière d'hétéroscédasticité, comme le GARCH par exemple.

Une sorte de révolution a été faite par Politis, Romano et Wolf (1996) lorsqu'ils ont utilisé la méthode de sous-échantillonnage (subsampling) pour trouver la fiabilité des intervalles de confiance des paramètres de régression dans un contexte de dépendance et de possibilité d'hétéroscédasticité dans les données. Dans ce cadre, la procédure de sous-échantillonnage a servi à générer des pseudo séries de rendements. La méthode de sous-échantillonnage est une approche très avantageuse quand le bootstrap ne peut pas être appliquée, pour augmenter la fiabilité des tests statistiques. Par contre, dans le cas contraire, il est préférable d'utiliser le bootstrap. Dans leur étude Politis, Romano et Wolf (1996) ont utilisé la statistique ratio de variance (variance ratio statistic) comme statistique de test. En considérant trois horizons $k=2,4$ et 8^5 , ils ont trouvé de l'évidence de prévisibilité pour le « CRSP equal-weighted index ». Tandis que, pour le « CRSP value-weighted index » et pour le « S&P 500 index » Politis, Romano et Wolf (1996) n'ont pas pu rejeter la nulle de non prévisibilité des rendements financiers.

Le même travail a été repris par Wolf (2000) où il a testé si le coefficient de régression des rendements continûment composés de k périodes sur le ratio dividende-prix est statistiquement différent de zéro. En utilisant la méthode de sous-échantillonnage sur trois séries de données, où la taille des blocs est choisie avec une technique de calibrage, Wolf (2000) n'a pas trouvé de preuve convaincante pour la prévisibilité des rendements des actifs financiers.

⁵ L'horizon k est mesuré en mois.

Remarquons que Wolf (2000) n'a utilisé le modèle VAR(1) avec deux variables que pour fin de simulation. Il a pu conclure sur la meilleure performance de la méthode de sous-échantillonnage par rapport à celle de GMM. Il est à préciser que, même si cette méthode fonctionne sans imposer trop de conditionnalité et sans avoir choisir le modèle structurel pour les données observées, le fait de choisir la taille du bloc d'une façon ad hoc la rend moins attrayante. Wolf (2000) dans son article essayait de surmonter ce problème en utilisant une technique de calibrage. Notons que même si la technique de calibrage permet de résoudre le problème du choix de la taille du bloc, elle reste une technique difficile à implémenter.

Goetzmann et Jorion (1993), quant à eux, ont utilisé une approche bootstrap pour générer des séries des données artificielles sous la nulle. Dans cette approche, les rendements sont considérés comme *i.i.d.* selon une distribution inconnue ce qui implique que les dividendes sont parfaitement non stochastiques. Ainsi, on peut dire que Goetzmann et Jorion (1993) ne tiennent pas compte de la présence d'hétéroscédasticité dans les rendements financiers. En suivant cette démarche, Goetzmann et Jorion (1993) n'ont pas trouvé une forte évidence statistique en faveur de la prévisibilité des rendements financiers.

Malliaropulos (1996) a utilisé aussi une méthode de bootstrap *i.i.d.* pour générer les pseudo séries de rendements. Il a fait les tests de prévisibilité à l'aide des ratios de variance (variance ratio statistics) univarié et multivarié. En utilisant des données mensuelles de FT.A All Share index de janvier 1964 à septembre 1993, il n'était pas capable de rejeter l'hypothèse que les log rendements sont *i.i.d.*⁶ Il est notoire que Malliaropulos (1996) a utilisé une procédure de bootstrap qui n'est en mesure de tester que l'hypothèse *i.i.d.*

⁶ Le fait que Malliaropulos (1996) a généré les pseudo rendements comme *i.i.d.* ne lui permet de tester que l'hypothèse forte de l'indépendance sérielle des log rendements.

D'autre part, Hodrick (1992) a considéré le problème d'hétéroscédasticité dans les données financières d'une façon paramétrique. Dès lors, il a utilisé un modèle VAR(1) avec trois variables dépendantes: le logarithme des rendements réels ($\log R_t$), le ratio dividende prix annualisé (D_t/P_t) et le rendement du bon de trésor relatif à sa moyenne mobile des douze derniers mois (rb_t).

Soit Z_t le vecteur des variables dépendantes du modèle VAR(1). Chaque variable est centrée par rapport à sa moyenne :

$$Z_t \equiv [\log R_t - E(\log R_t), D_t/P_t - E(D_t/P_t), rb_t - E(rb_t)]' \quad (11)$$

$$Z_{t+1} = AZ_t + \mu_{t+1} \quad (12)$$

où A est une matrice 3×3 avec élément typique a_{ij} et μ_t est un bruit blanc multivarié.

L'hypothèse nulle d'intérêt est

$$H_0 : a_{11} = a_{12} = a_{13} = 0 \quad (13)$$

Hodrick (1992) a comparé les distributions asymptotiques avec les distributions asymptotiques avec les distributions d'échantillonnage en simulant des données sous la nulle. Pour tenir compte de l'hétéroscédasticité, il a postulé un modèle paramétrique d'hétéroscédasticité conditionnelle pour le vecteur d'innovations: un GARCH multivarié. Pour ce modèle spécifique, il a conclu que les distributions asymptotiques été fiables à l'échantillon fini. De plus, Hodrick a trouvé des résultats favorables à la prévisibilité des rendements des actifs financiers. Remarquons, cependant, que ces conclusions dépendent du modèle paramétrique estimé par Hodrick (1992). Qu'arrive-t-il si les données sont hétéroscédastiques mais qui ne sont pas générées exactement d'après un GARCH ?

Une solution pour surmonter ce problème est d'utiliser le bootstrap. Mais, qu'est-ce que le bootstrap ?

2. Le bootstrap

La procédure du bootstrap a été introduite par Efron (1979). Elle est basée sur l'idée de rééchantillonnage des données. La méthode du bootstrap permet d'approximer la distribution d'un estimateur ou d'une statistique de test en rééchantillonnant des données. Il s'agit d'une méthode pour trouver des approximations de quantités qui peuvent être difficilement calculables.

L'idée est de remplacer la fonction de distribution du terme d'erreur laquelle est souvent la loi normale par la fonction de distribution des résidus, (voir Davidson (avr.98)). L'échantillon utilisé est considéré comme une population dans laquelle on fait des tirages; en rééchantillonnant, de nouveaux échantillons sont créés et en répétant cette procédure un certain nombre de fois, il est possible d'utiliser la moyenne de ces quantités calculées afin d'obtenir une estimation de la valeur espérée de cette quantité « bootstrappée ».

En vue d'illustrer la procédure du bootstrap, nous allons prendre un exemple.

2.1. Un exemple de bootstrap: le cas de la moyenne empirique

Soit $X = \{X_1, X_2, \dots, X_n\}$ un échantillon *i.i.d.* à partir d'une population F , c'est-à-dire.

$$F(x) = P(X_i \leq x), \quad i=1, \dots, n \quad (14)$$

La moyenne de la population $\theta(F) = E_F X_1$, est estimée par la moyenne échantillonnale

$$\hat{\theta}_n \equiv \hat{\theta}_n(X) = \bar{X}_n \quad (15)$$

Le but est d'obtenir un estimateur de la variance de $\hat{\theta}_n$ tel que

$$\text{var}_F(\hat{\theta}_n) = \frac{\sigma_F^2}{n} \quad (16)$$

avec $\sigma_F^2 = \text{var}_F(X_i)$

Notons que l'estimateur traditionnel de σ_F^2 est $\hat{\sigma}_F^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$

Supposons, pour le moment, que F est connue. Dans ce cas, σ_F^2 peut être approximée par une simulation de Monte Carlo. Nous pouvons générer B rééchantillons *i.i.d.* X^1, \dots, X^B , où chaque échantillon X^k est formé par n observations *i.i.d.* à partir de F .

Si B est assez grand, la Loi des Grands Nombres (LGN) stipule que :

$$E_F g(T(X)) \approx \frac{1}{B} \sum_{i=1}^B g(T(X^{(i)})) \quad (17)$$

où $g(\cdot)$ est une fonction quelconque, exemples, $g(x)=x$ ou $g(x)=x^2$

Ainsi, nous obtenons :

$$\text{var}_F(\hat{\theta}_n) \approx B^{-1} \sum_{k=1}^B \hat{\theta}_n^2(X^k) - \left[B^{-1} \sum_{k=1}^B \hat{\theta}_n^2(X^k) \right]^2 \quad (18)$$

La procédure du bootstrap consiste dans ce cas à:

1) construire un nouvel échantillon (X_1^*, \dots, X_n^*) à partir de l'échantillon original.

Comme, par exemple, pour obtenir un pseudo échantillon particulier X^* , nous pouvons utiliser un mécanisme de générateur de nombres aléatoires afin de tirer des entiers I_1, \dots, I_n où chacun d'eux est tiré entre $1, \dots, n$ avec une probabilité de $1/n$. Dans ce cas, notre rééchantillon est formé de $(X_{I_1}, \dots, X_{I_n})$.

2) calculer la moyenne empirique de chaque rééchantillon :

$$\hat{\theta}_n(X^*) = \frac{1}{n} \sum_{i=1}^n X_i^* \quad (19)$$

3) répéter les deux étapes précédentes B fois.

Le fait que la population, où les rééchantillons de bootstrap sont tirés, est connue nous permet de calculer la variance (population variance) de $\hat{\theta}_n^*$:

$$\begin{aligned} \text{var}^*(\hat{\theta}_n^*) &= \text{var}^*\left(n^{-1} \sum_{i=1}^n X_i^*\right) \\ &= n^{-1} E^*\left[\left(X_1^* - E^*(X_1^*)\right)^2\right] \\ &= n^{-1} \left[n^{-1} \sum_{i=1}^n \left(X_i - \bar{X}_n\right)^2 \right] \end{aligned} \quad (20)$$

La variance représentée ci-dessus est l'estimateur idéal de bootstrap de la variance, qui correspond à prendre $B = \infty$.

Malheureusement, dans les problèmes pratiques, la population et sa distribution F sont inconnues. C'est ainsi que le bootstrap arrive en tant que solution. L'idée de la méthode de bootstrap est de faire le mieux possible avec le peu d'informations que nous avons, qui est l'échantillon observé $X=(X_1, \dots, X_n)$. Cette phrase est une traduction libre de celle de Politis (1997) : « Since you do not have the whole population, do the best with you do have, which is the observed sample $X=(X_1, \dots, X_n)$ »

L'estimateur de la variance de bootstrap est donné par la variance empirique de $\hat{\theta}_n^*$ des B rééchantillons.

$$B^{-1} \sum_{k=1}^B \left(\hat{\theta}_n^{*k}\right)^2 - \left[B^{-1} \sum_{k=1}^B \hat{\theta}_n^{*k} \right]^2 \approx \text{var}^*(\hat{\theta}_n^*) \quad (21)$$

Remarquons que, comme $\text{var}^*(\hat{\theta}_n^*) \xrightarrow{P} \text{var}_{F^*}(\hat{\theta}_n)$, le bootstrap fonctionne.

Il est à noter que dans de nombreuses situations le bootstrap permet de réaliser des tests d'hypothèses plus fiables en petit échantillon que ceux fondés sur la théorie asymptotique. Une question se pose ici, comment effectuer un test basé sur le bootstrap?

2.2. Le test bootstrap classique

Le bootstrap est une méthode forte utilisatrice d'informatiques. Il substitue la puissance de l'ordinateur aux formules analytiques compliquées. L'idée du bootstrap est de générer un grand nombre d'échantillons bootstrap qui suit l'hypothèse nulle. Il est alors requis que la distribution de la statistique τ , statistique de test, soit différente sous la nulle et sous l'hypothèse alternative afin que la statistique puisse discriminer entre les deux hypothèses.⁷

Il est notoire que, pour que le test bootstrap fonctionne bien⁸, la statistique de test doit être un pivot asymptotique, c'est-à-dire que sa distribution asymptotique ne doit pas dépendre d'un quelconque élément du processus générateur de données (PGD). Autrement dit, asymptotiquement la distribution de la statistique de test est la même pour tout processus générateur de données.

Le processus générateur de données qui a généré les données originales est inconnu. Le PGD bootstrap est alors une estimation du vrai PGD inconnu. Si le PGD bootstrap est proche du vrai PGD, alors les données seront proches de celles générées par le

⁷ Cf. Mackinnon, 1999

⁸ Si la statistique de test τ est pivotale, alors sa distribution sous le PGD bootstrap, $\hat{\mu}$ est la même que celle sous le vrai PGD, μ . Par conséquent, $\Pr_{\hat{\mu}}(\tau \geq \hat{\tau}) = \Pr_{\mu}(\tau \geq \hat{\tau})$. Ainsi, dans ce cas, le test bootstrap idéal ($B \longrightarrow \infty$) fonctionne bien sous la nulle. Autrement dit, il permet de contrôler l'erreur de type 1.

vrai processus générateur de données. Dans ce cas, la p-value du PGD bootstrap devrait être proche de la vraie p-value.

La procédure de test usuelle du bootstrap, pour un test unilatéral où nous rejetons la nulle lorsque la valeur réalisée de la statistique de test, $\hat{\tau}$, est suffisamment grande, consiste en les étapes suivantes pour calculer une p-value bootstrap (Davidson R. (avr.98)) :

- 1) estimer le modèle de manière traditionnelle. Calculer la statistique de test $\hat{\tau}$, et les paramètres du modèle.
- 2) réaliser un grand nombre d'itérations (noté B) des quatre procédures suivantes:
 - i) générer n nombres (pseudo) aléatoires $u_j^* \sim N(0, 1)$
 - ii) recréer le processus générateur de données bootstrap sous H_0 pour chaque série, estimer le modèle à partir des paramètres estimés auparavant sous H_0
 - iii) estimer le modèle. Calculer la statistique de test $\hat{\tau}^*$
 - iv) calculer la fonction de p-value bootstrap telle que

$$p^*(\hat{\tau}) = \frac{\#(\hat{\tau}^* > \hat{\tau})}{B} \quad (22)$$

où $\#(\hat{\tau}^* > \hat{\tau})$ est une fonction indicatrice binaire: elle vaut 1 quand le test $\hat{\tau}^* > \hat{\tau}$ est vrai et 0 sinon.

- 3) Si $p^*(\hat{\tau}) < \alpha$, avec α le niveau du test, l'hypothèse nulle est rejetée.

Dans de nombreuses situations, il est préférable que le nombre de répétitions du test bootstrap tende vers l'infini.

Le test bootstrap présenté ci haut est fondé sur l'hypothèse que les résidus sont gaussiens. Cette hypothèse peut-être contraignante, notamment lorsqu'il s'agit de séries financières. Quel type de bootstrap pouvons-nous utiliser dans ce cas de figure?

2.3. Le wild bootstrap (WB) récursif

La méthode de rééchantillonnage bootstrap standard d'Efron repose sur le fait que les erreurs sont identiquement et indépendamment distribuées. Cette hypothèse n'est pas satisfaite dans la plupart des cas en séries temporelles, entre autres les séries financières, où le modèle pourrait présenter de l'hétéroscédasticité et/ou de l'autocorrélation.

Le processus générateur de données ne peut être reconstruit de manière traditionnelle; il doit l'être de manière récursive en supposant la valeur initiale de la série temporelle donnée. Cette méthode permet de conserver la structure de corrélation des erreurs.

Le wild bootstrap récursif⁹ est une simple modification de la méthode usuelle de bootstrap récursif pour autorégressions qui consiste à remplacer le *i.i.d* bootstrap d'Efron par le wild bootstrap au moment de « bootstrapper » les erreurs du modèle AR.¹⁰

D'une façon plus spécifique, le wild bootstrap récursif génère les pseudo séries, étant donné le processus autorégressif, de la façon suivante:

⁹ Il existe d'autres méthodes de WB (cf Gonçalves et Kilian, 2002) mais le WB récursif paraît le plus naturel.

¹⁰ Cf. Gonçalves et Kilian, 2002

$$y_t^* = \hat{\phi}_{1n} y_{t-1}^* + \hat{\phi}_{2n} y_{t-2}^* + \dots + \hat{\phi}_{pn} y_{t-p}^* + \hat{\varepsilon}_t^*, \quad t=1, \dots, n \quad (23)$$

où

$$\hat{\varepsilon}_t^* = \hat{\varepsilon}_t \eta_t \text{ avec } \hat{\varepsilon}_t = \hat{\phi}_n(L) y_t \quad (24)$$

où η_t est une suite *i.i.d.* de moyenne zéro et de variance égale à l'unité telle que

$$E^* |\eta_t|^4 \leq \Delta < \infty \text{ et } \hat{\phi}_n(L) = 1 - \hat{\phi}_{1n}L - \hat{\phi}_{2n}L^2 - \dots - \hat{\phi}_{pn}L^p, \quad \hat{\phi}_{pn} \neq 0 \quad (25)$$

La valeur initiale de y_t^* est donnée, par exemple, $y_0^* = 0$.

Notons aussi que les nouvelles innovations de WB récursif a pour moyenne zéro et variance égale à $\hat{\varepsilon}_t^2$ comme dans les cas des modèles ARCH, GARCH.

Dès lors,

$$\begin{aligned} E^* \left(\hat{\varepsilon}_t^* \right) &= E^* \left(\hat{\varepsilon}_t \eta_t \right) \\ E^* \left(\hat{\varepsilon}_t^* \right) &= \hat{\varepsilon}_t E^* \left(\eta_t \right), \text{ car } \hat{\varepsilon}_t \text{ est une fonction des données} \\ E^* \left(\hat{\varepsilon}_t^* \right) &= 0, \text{ par hypothèse } \eta_t \text{ a une moyenne nulle} \end{aligned} \quad (26)$$

De même,

$$\begin{aligned} \text{var}^* \left(\hat{\varepsilon}_t^* \right) &= \text{var}^* \left(\hat{\varepsilon}_t \eta_t \right) \\ \text{var}^* \left(\hat{\varepsilon}_t^* \right) &= \hat{\varepsilon}_t^2 \text{var}^* \left(\eta_t \right), \text{ car } \hat{\varepsilon}_t \text{ est une fonction des données} \\ \text{var}^* \left(\hat{\varepsilon}_t^* \right) &= \hat{\varepsilon}_t^2, \text{ par hypothèse } \eta_t \text{ a une variance égale à l'unité} \end{aligned} \quad (27)$$

Ce type de bootstrap est robuste à la présence d'hétéroscédasticité conditionnelle dans les erreurs si ceux-ci suivent un « martingale difference sequences ». ¹¹ La forme de l'hétéroscédasticité peut ne pas être connue.

Dans le cas de WB récursif, la procédure est la suivante :

1) estimer le modèle $AR(p)$ par MCO afin d'obtenir les résidus

$$\hat{\varepsilon}_t = \hat{\phi}_n(L) y_t, t=1, \dots, n \quad (28)$$

2) reconstruire les résidus selon le wild bootstrap

$$\hat{\varepsilon}_t^* = \hat{\varepsilon}_t \eta_t, \eta_t \sim i.i.d.(0,1) \quad (29)$$

3) générer les pseudo séries de la façon suivante: pour $t=1, \dots, n$, soit

$$y_t^* = \hat{\phi}_{1n} y_{t-1}^* + \hat{\phi}_{2n} y_{t-2}^* + \dots + \hat{\phi}_{pn} y_{t-p}^* + \hat{\varepsilon}_t^* \quad (30)$$

S'il en est ainsi la présentation de procédure du bootstrap, comment pouvons-nous l'utiliser pour faire notre test sur la prévisibilité des rendements d'actifs financiers?

¹¹ Cf. Gonçalves et Kilian, 2002

3. Le test de prévisibilité des rendements financiers par le WB récursif

Dans cette section, nous présenterons une manière de faire le test de prévisibilité des rendements financiers qui est robuste à la présence d'hétéroscédasticité conditionnelle dans les données. Nous allons d'abord décrire le modèle utilisé.

3.1. Le modèle

Dans le présent rapport, nous allons faire deux sortes de test de prévisibilité. Le premier est un test non paramétrique de ratio de variance tandis que le deuxième est un test basé sur le ratio de variance dérivé à partir d'un modèle VAR.

Concernant la deuxième approche, le modèle utilisé est un modèle VAR avec trois variables dépendantes, à savoir le logarithme des rendements réels, $\log R_t$, le ratio dividende-prix, D_t/P_t , et le rendement du bon de trésor d'un mois relatif à sa moyenne mobile des douze derniers mois, rb_t , comme celui de Hodrick (1992). L'ordre du VAR est choisi à partir du critère de Schwarz (1978)¹². L'utilisation du modèle VAR nous permet de prendre en considération l'autocorrélation existante dans les données. Le VAR caractérise complètement les autocovariances des séries. Il permet, par exemple, aux rendements passés d'influencer les dividendes.

Notons que d'une façon générale, un modèle VAR(p) se présente comme suit :

$$y_t = v + A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t, \quad t = 0, \pm 1, \pm 2, \dots \quad (31)$$

où $y_t = (y_{1t}, \dots, y_{kt})'$ est un vecteur aléatoire de dimension $(k \times 1)$, les A_i sont des matrices fixes $(k \times k)$, $v = (v_1, \dots, v_k)'$ est le vecteur des constantes de dimension

¹² Cf. Annexe I

$(k \times 1)$ ¹³. Par ailleurs, $u_t = (u_{1t}, \dots, u_{kt})'$ est un bruit blanc de dimension k , tel que, $E(u_t) = 0$, $E(u_t u_t') = \Sigma_u$, et $E(u_t u_s') = 0$ pour $s \neq t$

La matrice de variance-covariance Σ_u est assumée non singulière.

Le modèle $VAR(p)$ décrit ci-haut peut-être écrit sous forme compacte à un modèle $VAR(1)$ de dimension kp .

$$Y_t = v + AY_{t-1} + U_t \quad (32)$$

où

$$Y_t := \begin{bmatrix} y_t \\ y_{t-1} \\ \cdot \\ \cdot \\ y_{t-p+1} \end{bmatrix} \quad (kp \times 1) \quad v := \begin{bmatrix} v \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \quad (kp \times 1) \quad A := \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I_k & 0 & \dots & 0 & 0 \\ 0 & I_k & \dots & 0 & 0 \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & \dots & I_k & 0 \end{bmatrix} \quad (kp \times kp) \quad U_t := \begin{bmatrix} u_t \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \quad (kp \times 1)$$

Chaque test statistique est basé sur une statistique qu'on appelle statistique de test. Celle-ci sert à discriminer entre les deux hypothèses nulle et alternative. Il s'avère alors nécessaire de présenter la statistique de test utilisée pour notre test de prévisibilité des rendements financiers.

¹³ Celui-ci permet à $E(y_t)$ d'être différente de zéro.

3.2. Le ratio de variance (variance ratio statistics)

Une version stylisée de l'hypothèse de marché efficient stipule que les rendements d'actifs financiers sont indépendants sériellement et donc non prévisibles. Une hypothèse plus faible et plus intéressante est que le logarithme des rendements sont non autocorrélés, mais peuvent être dépendants. Par ailleurs, plusieurs récents travaux soutiennent que les rendements d'actifs financiers peuvent être partiellement prévisibles, soit par les rendements passés ou autres variables comme le ratio dividende-prix. Une façon de tester la prévisibilité des rendements à partir des rendements passés est d'utiliser le test ratio de variance (variance ratio test). Ce test utilise le simple fait que la variance de la somme des variables aléatoires non corrélées est égale à la somme des variances individuelles (Politis, Romano et Wolf, 1997).

La simulation de Monte Carlo faite par Poterba et Summers (1988) confirme que les tests par le ratio de variance de type utilisés par Lo et Mackinlay (1988) et Kim, Nelson et Startz (1991) sont les plus puissants pour tester l'efficacité des marchés financiers au sens faible.

Dans notre travail, nous allons distinguer deux sortes de ratio de variance. L'un est univarié¹⁴ et l'autre est multivarié dérivé à partir du modèle VAR. Par ailleurs, le test par le ratio de variance multivarié est plus intéressant, que celui de l'univarié, dans la mesure où le modèle VAR permet d'utiliser un plus grand ensemble d'informations pour prédire les rendements financiers.

¹⁴ Il est à remarquer que le ratio de variance univarié nous permet de tester l'hypothèse de marche aléatoire de type 3 (RW3) où les rendements peuvent être dépendants mais non corrélés. Ce test est une manière de tester la version faible de l'hypothèse de l'efficacité des marchés des actifs.

3.2.1. Le ratio de variance univarié

Admettons que nous observons le prix de l'actif - financier - au même intervalle de temps i (ex : chaque mois).

Soit le rendement réel total d'une période

$$R_{i+1} = \frac{P_{i+1} + d_{i+1}}{P_i} \quad (33)$$

où P_i est le prix réel de l'actif à la fin de la période i et d_i est le rendement réel en dividende payé durant la période i . Notons que d_i peut être nul dans le cas où l'actif ne paie pas de dividende.

Admettons aussi que le log rendements $r_i \equiv \log(R_i)$ est stationnaire. Nous sommes intéressés à tester l'hypothèse nulle que les rendements sont non corrélés :

H_0 : $r_i \equiv \log(R_i)$ sont non autocorrélés avec une moyenne μ et une variance σ^2

Soit un échantillon de mk observations, où k est un entier supérieur ou égal à l'unité, r_1, \dots, r_{mk} . Les estimateurs sans biais de μ et σ^2 sont donnés par :

$$\begin{aligned} \hat{\mu} &= mk^{-1} \sum_{i=1}^{mk} r_i \\ \hat{\sigma}_1^2 &= \frac{1}{mk-1} \sum_{i=1}^{mk} (r_i - \hat{\mu})^2 \end{aligned} \quad (34)$$

Sous H_0 , $\hat{\sigma}_k^2$ est aussi un estimateur sans biais de σ^2 tel que :

$$\hat{\sigma}_k^2 = \frac{1}{l} \sum_{i=k}^{mk} \left(p_i - p_{i-k} - k \hat{\mu} \right)^2 \quad (35)$$

où $l = k(mk - k + 1) \left(1 - \frac{k}{mk} \right)$

avec $p_i \equiv \log(P_i)$

Remarquons que $ri \equiv p_i - p_{i-1}$

$$\text{Ainsi, } p_i - p_{i-k} = \sum_{j=1}^k r_{i+j-1}$$

$$\text{D'où } \hat{\sigma}_k^2 = \frac{1}{l} \sum_{i=k}^{mk} \left(\sum_{j=1}^k r_{i+j-1} - k \hat{\mu} \right)^2$$

Cependant, si les log rendements sont autocorrélés, $\hat{\sigma}_k^2$ ne serait plus un estimateur sans biais de σ^2 . Sous une autocorrélation positive, $\hat{\sigma}_k^2$ tend à être plus grand que σ^2 .

Tandis que sous une autocorrélation négative $\hat{\sigma}_k^2$ tend à être plus petit que σ^2 .

La statistique ratio de variance (variance ratio test statistic) pour un horizon k peut être alors définie par :

$$M_r(k) = \frac{\hat{\sigma}_k^2}{\hat{\sigma}_1^2} \quad (36)$$

Signalons que $M_r(k)$ peut être asymptotiquement approximée par :

$$M_r(k) \approx 1 + 2 \sum_{j=1}^{k-1} \left(1 - \frac{j}{k}\right) \hat{\rho}_j \quad (37)$$

où $\hat{\rho}_j$ est le coefficient d'autocorrélation échantillonnale d'ordre k de $\{r_i\}$.

$M_r(k)$ est un estimateur convergent de la statistique ratio de variance pour un horizon k :

$$VR(k) = 1 + 2 \sum_{j=1}^{k-1} \left(1 - \frac{j}{k}\right) \rho_j \quad (38)$$

où ρ_j est le coefficient d'autocorrélation d'ordre k de $\{r_i\}$.

La statistique $M_r(k)$ est celle utilisée par Lo et Mackinlay (1988). Il est évident que sous la nulle, $M_r(k)$ tend vers l'unité. Pourtant, si les log rendements r_i sont

autocorrélés $M_r(k)$ tend à être plus grand que l'unité, et vice-versa pour une autocorrélation négative.

Regardons maintenant la statistique ratio de variance dérivée à partir du VAR.

3.2.2. Le ratio de variance multivarié

Sans perte de généralité, assumons que le modèle VAR retenu, pour représenter les données, est d'ordre un.¹⁵

Soit Z_t le vecteur des variables dépendantes du modèle VAR(1). Chaque variable est centrée par rapport à sa moyenne :

$$Z_t \equiv [\log R_t - E(\log R_t), D_t / P_t - E(D_t / P_t), rb_t - E(rb_t)]' \quad (39)$$

$$Z_{t+1} = AZ_t + \mu_{t+1} \quad (40)$$

où A est une matrice 3×3 avec élément typique a_{ij} et μ_t est un bruit blanc multivarié.

Notons que $E(Z_{t+i} | \Phi_t) = A^i Z_t$,

où Φ_t est l'ensemble des informations disponibles avant et à la date t .

Si les séries sont stationnaires, l'équation (40) implique que

$$\begin{aligned} Z_{t+1} &= (I - AL)^{-1} u_{t+1} \\ &= \sum_{j=0}^{\infty} A^j u_{t+1-j} \end{aligned} \quad (41)$$

¹⁵ Celle-ci n'est pas contraignante. En effet, nous avons vu dans la sous section précédente que tout modèle VAR peut-être représenté par un modèle VAR(1).

Ainsi, la variance inconditionnelle de Z_t est

$$\begin{aligned}
 C(0) &= \text{var} \left(\sum_{j=0}^{\infty} A^j u_{t+1-j} \right) \\
 &= \sum_{j=0}^{\infty} A^j \text{var}(u_{t+1-j}) A^{j'} \\
 &= \sum_{j=0}^{\infty} A^j V A^{j'}
 \end{aligned} \tag{42}$$

avec $V = \text{var}(u_{t+1-j})$

Pour tenir compte des rendements composés sur k périodes, considérons la somme de k Z_t consécutives dont k représente l'horizon de la prévision. La variance de cette somme est :

$$\begin{aligned}
 V_k &\equiv \text{var} \left(\sum_{j=0}^{\infty} Z_{t+j} \right) \\
 &= kC(0) + \sum_{j=1}^{k-1} (k-j) [C(j) + C(j)']
 \end{aligned} \tag{43}$$

avec $C(j) = A^j C(0)$ ¹⁶

La variance totale de la somme des k rendements est l'élément (1,1) de V_k . Elle est donnée par $e_1' V_k e_1$ où $e_1 = (1, 0, 0)'$

Ainsi, la statistique ratio de variance (variance ratio), à partir du VAR, peut être écrite comme :

$$VR(k) = \frac{e_1' V_k e_1}{k e_1' C(0) e_1} \tag{44}$$

¹⁶ Nous avons tronqué la somme infinie de $C(0)$ à 127 comme dans Hodrick (1992).

Notons que si les rendements sont non corrélés, c'est-à-dire non prévisibles, la variance de la somme des k rendements sera égale à k fois la variance d'un rendement. Ainsi, le ratio de variance, $VR(k)$, sera égale à l'unité. $VR(k)$ peut être inférieure ou bien supérieure à l'unité dans les cas où les rendements sont corrélés négativement ou bien positivement.

Nous sommes, maintenant, en mesure de décrire notre test de prévisibilité des rendements financiers.

3.3. La démarche du test

Notons que pour estimer la distribution échantillonnale de la statistique de test, des pseudo séries sont générées à partir de leurs distributions empiriques en utilisant le WB récursif.

La procédure de test suivie consiste en les étapes suivantes pour construire un intervalle de confiance de bootstrap pour la statistique $M_r(k)$:

- 1) Générer une pseudo série de log rendements $\{r_i^*\}$ telle que $r_i^* = r_i \eta_i$
où $\eta_i \sim i.i.d(0,1)$ (ex : $\eta_i \sim N(0,1)$).
- 2) Calculer la statistique $M_r^*(k)$ à partir de $\{r_i^*\}$
- 3) Répéter 1) et 2) un certain nombre de fois, ex : B fois.
- 4) Ordonner les $M_r^*(k)$ et construire l'intervalle de confiance de bootstrap (equal-tailed $(1-\alpha)$ 100% bootstrap confidence interval) tel que :

$$IC_{M_r} = [M_{r_{k_1}}^*, M_{r_{k_2}}^*] \quad (45)$$

où k_1 est définie par $[B\alpha/2]+1$, et $k_2 : k_2 = [B(1-\alpha/2)]+1$. et $[.]$ représente la partie entière.

De même, la démarche de construction d'un intervalle de confiance de bootstrap pour $VR(k)$ peut être résumée comme suit :

- 1) Estimer par MCO le modèle VAR
- 2) Calculer la statistique $VR(k)$
- 3) Réaliser un grand nombre d'itérations (noté B) les procédures suivantes :
 - a) Obtenir les résidus d'après 1) et simuler une variable aléatoire externe, par exemple, $v_t \sim^{i.i.d.} N(0,1) \quad t=1, \dots, n$ ¹⁷ afin de générer les pseudo résidus de WB tels que $\hat{u}_t^* = \hat{u}_t v_t, \quad t=1, \dots, n$
 - b) Choisir une condition initiale. Par exemple, prendre au hasard une condition initiale parmi les données originales.
 - c) Utiliser le VAR estimé en 1) comme PGD après avoir imposé la nulle de non prévisibilité, c-à-d créer les pseudo séries Z_t^* .
 - d) Estimer le modèle par MCO. Calculer la statistique de test $VR^*(k)$
- 4) Construire l'intervalle de confiance de bootstrap telle que

$$IC_{VR} = [VR_{k_1}^*, VR_{k_2}^*] \quad (46)$$
 où k_1 est définie par $[B\alpha/2]+1$ et $k_2 : k_2 = [B(1-\alpha/2)]+1$, et $[.]$ représente la partie entière.
- 5) Si $1 \in IC_{VR}$, avec α le niveau du test, l'hypothèse nulle n'est pas rejetée.

¹⁷ D'après l'étude de Monte Carlo, le choix de la variable externe n'est pas importante lors du processus de WB récursif (cf Gonçalves et Kilian, 2002).

Notons que les intervalles de confiance décrits ci-dessus sont des intervalles qui tiennent compte de l'asymétrie qui peut exister dans la distribution de nos séries (voir Dimitris N. Politis, 1998)¹⁸. On les appelle en anglais : the equal-tailed $(1-\alpha)$ 100% bootstrap confidence interval.

S'il en est ainsi la présentation théorique de notre approche, passons maintenant à l'analyse empirique.

¹⁸ Il existe plusieurs façons de construire l'intervalle de confiance bootstrap. Dimitris N. Politis (1998) a constaté que "the equal-tailed $(1-\alpha)$ 100% bootstrap confidence interval" est l'un des plus populaires en cas de présence d'asymétrie.

4. Analyse empirique

Les données sont au centre de toute étude empirique. De ce fait, la description de celles-ci est primordiale avant toute analyse

4.1. Description des données

Les données sont du « Center for Research in Security Prices » (CRSP) du Centre interuniversitaire de recherche en analyse des organisations (CIRANO). Les séries sont mensuelles entre autres la valeur pondérée du rendement nominal avec dividende de NYSE, RN_t , la valeur pondérée du rendement nominal sans dividende, RX_t , le rendement du bon de trésor d'un mois, i_t , et le taux d'inflation de l'IPC, π_t . La période couverte s'étend de janvier 1926 à décembre 2001 soit 912 observations.

Pour avoir les séries de log rendements, $\ln R_t$, et le ratio dividende-prix, D_t/P_t , variables utilisées dans ce présent rapport, nous utilisons les transformations décrites dans l'appendice A, p. 382 et 383 du papier de Hodrick (1992). Par ailleurs, la série du rendement du bon de trésor relatif à sa moyenne mobile d'ordre douze, rb_t , est celle suggérée par Campbell (1991).

Ainsi,

$$rb_t = i_t - \frac{1}{12} \left(\sum_{j=1}^{12} i_{t-j} \right) \quad (47)$$

Notons qu'à cause des retards, les données sont réduites à 900 observations.

Le tableau suivant représente les statistiques descriptives de $\ln R_t$, D_t/P_t et rb_t ,

Tableau 1 : Statistiques descriptives des variables
(Données mensuelles de janvier 1927 et décembre 2001)

Séries	$\ln R_t$	D_t/P_t	rb_t
Nombre d'observations	900	900	900
Moyenne	.0056365	.0418982	-.0000107
Écart-type	.0545130	.0148742	.0103387
Minimum	-.3355548	.0149359	-.060534
Maximum	.3240505	.1340497	.068458

Nous remarquons que les données sont très variables. Cela se traduit par le fait que les écart-types des trois séries sont très élevés par rapport à leurs moyennes.

4.2. Résultats

Nous présenterons ci-dessous les résultats empiriques de notre travail. Nous allons commencer par présenter les résultats du test sur la prévisibilité des rendements financiers avec le ratio de variance univarié comme statistique de test.

Les intervalles de confiance asymétriques de bootstrap de la statistique ratio de variance univarié, $M_r(k)$, sont présentés par les Tableaux 2. Une valeur maximale de l'horizon k égale à 60 est retenue, correspondant à un horizon de rendement maximum de 5 ans. Ces intervalles sont calculés par la démarche décrite ci-dessus avec mille simulations de bootstrap. Trois échantillons sont retenus. Le premier couvre toutes nos observations de janvier 1927 à décembre 2001. Le second correspond à l'échantillon B (1952:01 à 1987:11) de Hodrick (1992) qui est l'échantillon choisi par celui-ci pour faire toute son analyse. Tandis que, le troisième part de décembre 1947 à décembre 1986. Ce dernier est l'échantillon considéré par Politis, Romano et Wolf (1996).

Tableaux 2 : « Equal-tailed 95% percentile bootstrap confidence interval » pour $M_r(k)$

Échantillon : janvier 1927 à décembre 2001

k	3	6	12	24	36	48	60
$M_r(k)$	1.1176	1.0691	1.1715	1.1223	1.0011	0.9265	0.8605
IC^{boot}	[0.84,1.17]	[0.77,1.29]	[0.68,1.45]	[0.60,1.73]	[0.54,1.84]	[0.46,1.98]	[0.46,2.14]

Échantillon : janvier 1952 à novembre 1987

k	3	6	12	24	36	48	60
$M_r(k)$	1.0817	1.1939	1.2593	1.2245	1.1081	1.0392	1.2042
IC^{boot}	[0.83,1.16]	[0.69,1.24]	[0.59,1.39]	[0.50,1.67]	[0.44,1.81]	[0.36,1.93]	[0.35,2.02]

Échantillon : décembre 1947 à décembre 1986

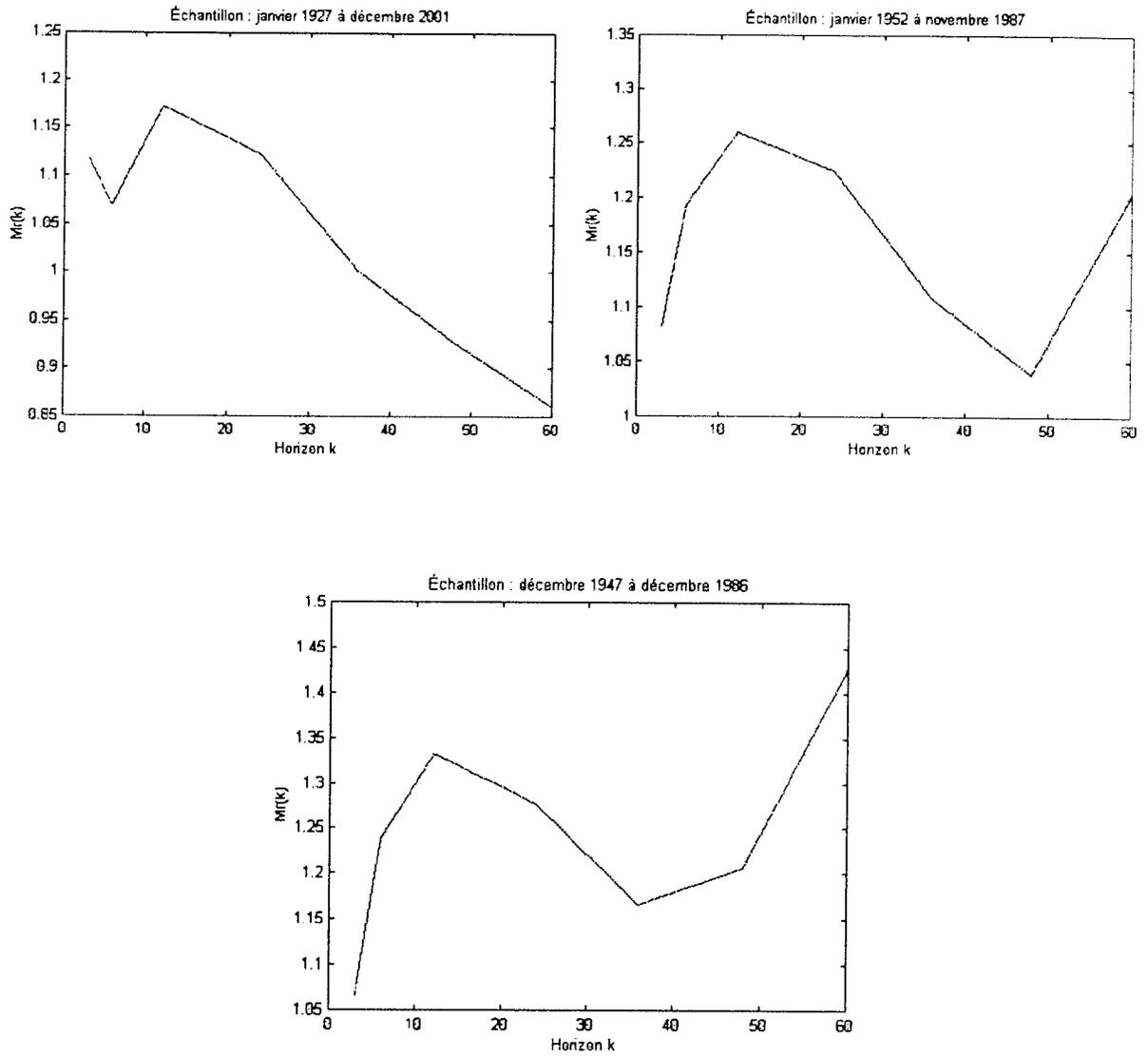
k	3	6	12	24	36	48	60
$M_r(k)$	1.0661	1.2389	1.3329	1.2771	1.1656	1.2055	1.4270
IC^{boot}	[0.85,1.16]	[0.79,1.28]	[0.65,1.44]	[0.54,1.62]	[0.46,1.75]	[0.43,2.02]	[0.38,2.16]

Les chiffres entre crochets représentent les « Equal-tailed 95% percentile bootstrap confidence interval » pour $M_r(k)$. Chaque IC^{boot} est obtenue à travers mille répliques de wild bootstrap récursif. Les ratios de variance $M_r(k)$ sont obtenus en utilisant les formules données en haut.

Les ratios de variance semblent, dans l'ensemble, augmenter avec l'horizon et descendre par la suite après que k dépasse 12 mois. Par ailleurs, pour les échantillons plus courts, les ratios de variance ont tendance à s'accroître quand l'horizon dépasse 48 mois. Les intervalles de confiance, quant à eux, deviennent de plus en plus larges avec l'horizon et ce, quelque soit l'échantillon considéré. De plus, ils sont assez

stables par rapport à l'échantillon considéré. Autrement dit, le choix de l'échantillon n'affecte pas beaucoup les intervalles de confiance de $M_r(k)$. Enfin, ils contiennent tous la valeur 1. L'hypothèse nulle de non corrélation des rendements financiers ne peut pas ainsi être rejetée. Cette conclusion n'est pas incompatible avec celle trouvée par Malliaropulos (1996) et Politis, Romano et Wolf (1996). Remarquons que ces deux papiers ont tous utilisé la statistique ratio de variance univarié comme statistique de test. Malliaropulos (1996) a choisi une procédure de bootstrap considérant les rendements comme *i.i.d.* Tandis que Politis, Romano et Wolf (1996) ont utilisé la méthode de sous échantillonnage. Notre approche, quant à elle génère les (pseudo) rendements de bootstrap par le wild bootstrap récursif.

L'approche de Politis, Romano et Wolf (1996) pour tester la prévisibilité des rendements financiers utilise une régression de long terme. Une telle régression fait face à deux types de problème : le problème d'hétéroscédasticité des rendements et celui de leur autocorrélation. Une manière appropriée, dans ce cas précis, de faire l'inférence sur la prévisibilité est d'utiliser la méthode de subsampling comme celle utilisée par Politis, Romano et Wolf (1996). Notons, par ailleurs, que même si la technique de calibrage surmonte le problème crucial du choix de la taille du bloc, elle reste difficile à implémenter. Ce caractère difficile rend l'approche de subsampling moins attrayante versus de celle de WB récursif. De plus, la performance relative du subsampling par rapport à la théorie asymptotique, en présence d'échantillons finis, n'a aucune justification théorique. Ce qui n'est pas le cas pour l'approche de bootstrap telle que le WB récursif.

Figure 1: Ratios de variance de log rendements pour l'horizon k , $M_r(k)$ 

Bref, le fait d'admettre que les données financières sont hétéroscédastiques, dont la forme de l'hétéroscédasticité peut-être inconnue, lors de l'étude sur la prévisibilité des rendements financiers nous a permis de ne pas rejeter l'hypothèse nulle d'absence d'autocorrélation dans la série des log rendements. De ce fait, les log rendements sont imprévisibles à partir de ses valeurs passées.

Nous allons maintenant tester la prévisibilité des rendements réels en employant une approche multivariée. Pour cette fin, un modèle VAR avec trois variables : le log rendements, le ratio dividende-prix et le rendement du bon de Trésor d'un mois relatif à sa moyenne mobile des douze derniers mois, est utilisé pour dériver la statistique ratio de variance multivarié. Comme dans le cas univarié, l'horizon maximum est fixé à cinq ans. La statistique est estimée en utilisant une simulation de bootstrap avec 1000 répétitions.

Le critère de Schwarz nous suggère, d'une part, un modèle VAR(2) pour les échantillons de janvier 1927 à novembre 1987 (échantillon A), de février 1927 à décembre 1951 (échantillon C), et de janvier 1927 à décembre 2001 (échantillon D) et, d'autre part, un modèle VAR(1) pour l'échantillon de janvier 1952 à novembre 1987 (échantillon B)¹⁹.

Les tableaux 3 reportent les résultats de l'estimation du modèle VAR avec différents échantillons.

¹⁹ Cf. Annexe II

Tableaux 3 : Estimation du modèle VAR de log rendements, de ratio dividende prix et de rendement relatif de bons de trésor

A : 1927:3 à 1987:11 (729 observations)

	$\ln(R_t)$	D_t/P_t	rb_t
$\ln(R_{t-1})$	-0.030627 (0.08613)	0.001243 (0.00043)	-0.001458 (0.00099)
$\ln(R_{t-2})$	-0.013805 (0.03844)	0.000203 (0.00019)	0.000496 (0.00044)
D_{t-1}/P_{t-1}	-28.57383 (16.8828)	1.381674 (0.08480)	-0.214272 (0.19326)
D_{t-2}/P_{t-2}	32.46211 (16.6721)	-0.415540 (0.08374)	0.172039 (0.19085)
rb_{t-1}	0.556950 (3.27047)	-0.023798 (0.01643)	0.571723 (0.03744)
rb_{t-2}	-4.514625 (3.27047)	0.023798 (0.01632)	0.127659 (0.03720)
Constante	-0.114721 (0.09148)	0.001441 (0.00046)	0.002019 (0.00105)
R^2	0.028235	0.938537	0.451181

B : 1952 :1 à 1987 :11 (431 observations)

	$\ln(R_t)$	D_t/P_t	rb_t
$\ln(R_{t-1})$	0.074246 (0.04957)	-0.000223 (0.00018)	-0.001374 (0.00094)
D_{t-1}/P_{t-1}	6.065545 (2.62833)	0.000203 (0.00019)	0.000496 (0.00044)
rb_{t-1}	-28.57383 (16.8828)	0.976957 (0.00977)	-0.118421 (0.04967)
Constante	-0.183646 (0.10881)	0.000905 (0.00040)	0.005028 (0.00206)
R^2	0.035472	0.959577	0.459538

C : 1927:3 à 1951:12 (298 observations)

	$\ln(R_t)$	D_t/P_t	rb_t
$\ln(R_{t-1})$	-0.152450 (0.14343)	0.001766 (0.00079)	-0.000716 (0.00067)
$\ln(R_{t-2})$	-0.009619 (0.05989)	0.000144 (0.00033)	0.000424 (0.00028)
D_{t-1}/P_{t-1}	-49.67469 (25.5094)	1.465354 (0.14114)	-0.113205 (0.11918)
D_{t-2}/P_{t-2}	54.40528 (24.8453)	-0.519686 (0.13746)	0.104655 (0.11608)
rb_{t-1}	14.02910 (11.9162)	-0.029695 (0.06593)	0.391916 (0.05567)
rb_{t-2}	3.407603 (12.0247)	0.010429 (0.06653)	0.303831 (0.05618)
Constante	-0.179898 (0.19326)	0.002761 (0.00107)	0.000362 (0.00090)
R^2	0.042737	0.906261	0.382612

D : 1927:3 à 2001:12 (898 observations)

	$\ln(R_t)$	D_t/P_t	rb_t
$\ln(R_{t-1})$	-0.047078 (0.07304)	0.001166 (0.00035)	-0.000990 (0.00084)
$\ln(R_{t-2})$	-0.027748 (0.03390)	0.000244 (0.00016)	0.000734 (0.00039)
D_{t-1}/P_{t-1}	-31.67756 (14.8685)	1.373541 (0.07203)	-0.124097 (0.17137)
D_{t-2}/P_{t-2}	34.12876 (14.1446)	-0.396398 (0.07143)	0.105849 (0.16994)
rb_{t-1}	-0.043139 (2.88829)	-0.002688 (0.01399)	0.589208 (0.03329)
rb_{t-2}	-3.169870 (2.88375)	0.019662 (0.01397)	0.141548 (0.03324)
Constante	-0.031951 (0.06724)	0.000841 (0.00033)	0.000755 (0.00077)
R^2	0.021872	0.955603	0.480186

Les variables sont constituées des rendements réels des actifs continûment composés, $\ln(R_t)$, les rendements en dividendes annualisés, D_t/P_t , et le rendement de bons de Trésor d'un mois relatif à sa moyenne mobile de douze derniers mois, rb_t . Les coefficients sont estimés par MCO. Les chiffres entre parenthèses représentent les écart-types estimés.

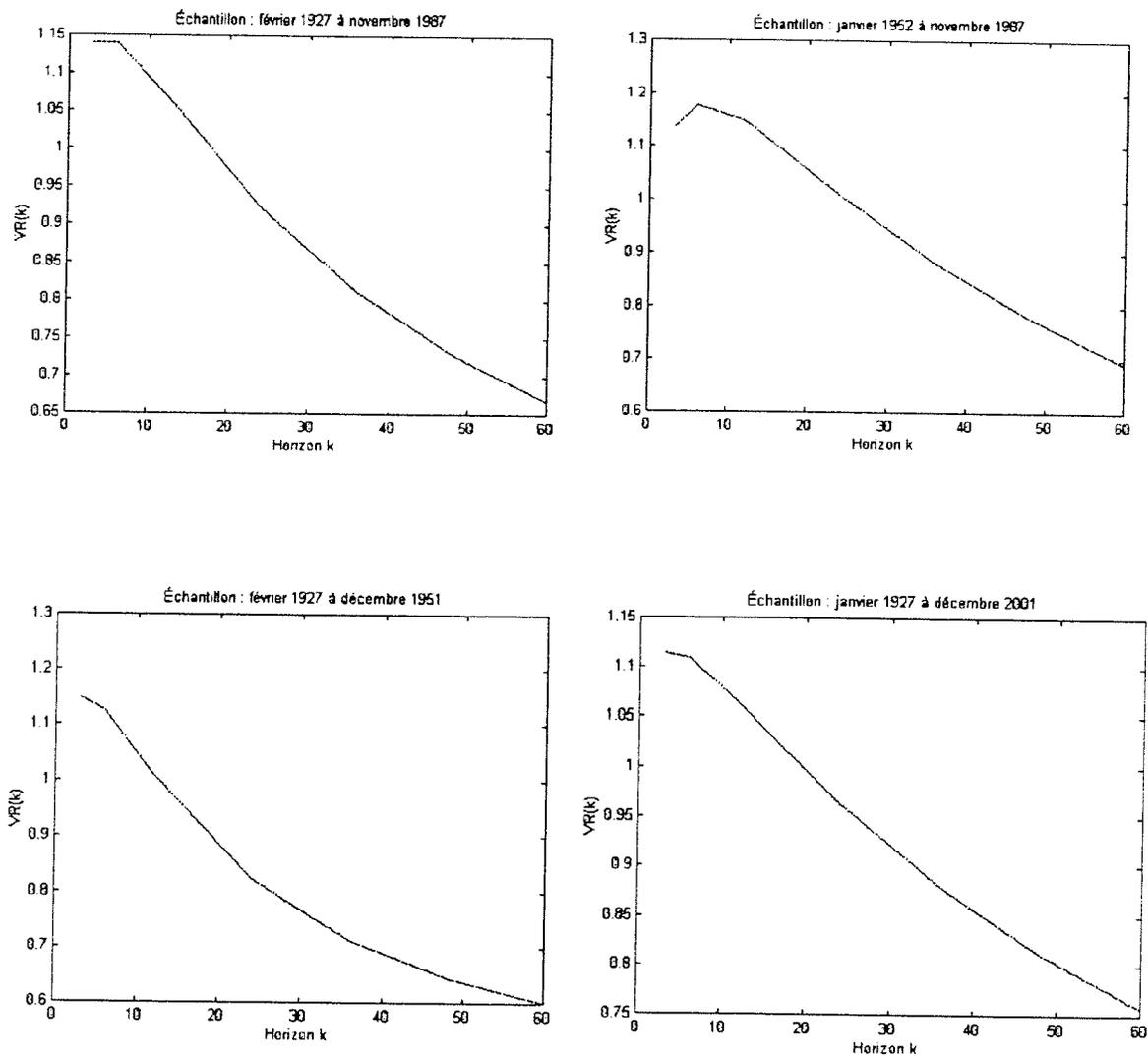
Notons que les estimateurs ponctuels du VAR ci haut sont utilisés pour générer les données artificielles pour la procédure du bootstrap. Les trois premiers échantillons sont ceux de Hodrick (1992) tandis que l'échantillon D recouvre toute la période observée (janvier 1927 à décembre 2001).

Le test de non prévisibilité revient à tester si les coefficients des variables retardées dans l'équation de log rendements sont nuls. Nous avons effectué ce test à l'aide d'un test de bootstrap robuste à l'hétéroscédasticité conditionnelle. La statistique de test est celle du ratio de variance. Nous présenterons ci-après les résultats de ce test.

Les tableaux 4 présentent les intervalles de confiance approximés par le bootstrap de la valeur de la statistique ratio de variance dérivée à partir du modèle VAR sous la nulle selon les différents échantillons. L'hypothèse nulle d'intérêt est l'absence d'autocorrélation entre les rendements réels dont aucune relation n'existe entre ces derniers et les autres variables du VAR. Les statistiques observées augmentent au dessus de l'unité pour les horizons de court terme (moins d'un an)²⁰ et diminuent petit à petit pour atteindre des valeurs proches de 0.7 à un horizon de 5 ans. Les intervalles de confiance, quant à eux, s'élargissent avec l'horizon peu importe l'échantillon. Ils restent stables par rapport à l'échantillon considéré. Notons que, comme ces intervalles contiennent tous la valeur unité, l'hypothèse nulle ne peut pas être rejetée. Ceci confirme les résultats trouvés avec l'approche univariée.

²⁰ À l'exception de l'échantillon C où les $VR(k)$ établissent une relation négative avec l'horizon k même à court terme.

Figure 2: Ratios de variance pour l'horizon k dérivés à partir du modèle VAR, $VR(k)$



Remarquons que Hodrick (1992) et Malliaropulos (1996) ont utilisé la même approche pour tester la prévisibilité des rendements financiers. Hodrick (1992) trouve des résultats favorables à la prévisibilité des rendements. Une conclusion différente à la nôtre. Une telle différence pourrait être due au fait que le modèle GARCH postulé par Hodrick (1992) n'est pas le vrai modèle. Malliaropulos (1996), quant à lui, n'a pas trouvé de l'évidence favorable à la prévisibilité des rendements financiers.

Tableaux 4 : « Equal-tailed 95% percentile bootstrap confidence interval » pour $VR(k)$ Échantillon : février 1927 à novembre 1987

Échantillon : mars 1927 à novembre 1987

k	3	6	12	24	36	48	60
$VR(k)$	1.1401	1.1404	1.0707	0.9236	0.8111	0.7288	0.6681
IC^{boot}	[0.84,1.16]	[0.78,1.26]	[0.70,1.41]	[0.61,1.62]	[0.55,1.86]	[0.51,2.01]	[0.51,1.89]

Échantillon : janvier 1952 à novembre 1987

k	3	6	12	24	36	48	60
$VR(k)$	1.1352	1.1774	1.1471	1.0114	0.8814	0.7785	0.6923
IC^{boot}	[0.84,1.16]	[0.79,1.20]	[0.70,1.28]	[0.57,1.42]	[0.48,1.64]	[0.43,1.80]	[0.41,1.81]

Échantillon : mars 1927 à décembre 1951

k	3	6	12	24	36	48	60
$VR(k)$	1.1473	1.1263	1.0085	0.8233	0.7126	0.6446	0.6004
IC^{boot}	[0.78,1.25]	[0.70,1.40]	[0.59,1.60]	[0.50,1.88]	[0.46,2.04]	[0.44,2.12]	[0.43,2.27]

Échantillon : mars 1927 à décembre 2001

k	3	6	12	24	36	48	60
$VR(k)$	1.1139	1.1093	1.0631	0.9636	0.8796	0.8117	0.7569
IC^{boot}	[0.85,1.15]	[0.80,1.24]	[0.73,1.34]	[0.65,1.50]	[0.60,1.60]	[0.56,1.74]	[0.52,1.94]

Les chiffres entre crochets représentent les « Equal-tailed 95% percentile bootstrap confidence interval » pour $VR(k)$. Chaque IC^{boot} est obtenue à travers mille répliquions de wild bootstrap récursif. Les ratios de variance $VR(k)$ sont obtenus en utilisant les formules données en haut.

Bref, la considération de plus d'informations via le modèle VAR ne change pas les conclusions trouvées avec l'approche univariée. Les rendements des actifs financiers restent imprévisibles aussi bien à court terme qu'à long terme.

CONCLUSION

Nous avons cherché à tester, dans ce travail, la prévisibilité des rendements financiers. Une question qui est devenue le leitmotiv des plusieurs études en économie financière. Dans un premier temps, nous avons passé en revue la littérature traitant la prévisibilité des rendements financiers. Nous avons essayé de mettre en évidence l'importance de la question, en particulier, dans la gestion optimale des portefeuilles. La prévisibilité des rendements financiers permet aux agents économiques de réallouer leurs portefeuilles de façon à maximiser leurs rendements espérés.

Étant conscients des rôles joués par la prévisibilité des rendements, les chercheurs s'efforçaient à trouver des relations liant les variables macroéconomiques et/ou financières aux rendements futurs espérés afin d'y pouvoir tester par la suite. Malheureusement, les tests sur la prévisibilité des rendements financiers ont tendance dans la plupart des cas soit à ignorer le fait empirique que les données financières sont hétéroscédastiques, soit à y imposer une forme particulière. Ces faits pourraient affecter les résultats des tests effectués. Dans ce rapport, nous avons utilisé une méthode robuste à l'hétéroscédasticité conditionnelle : le wild bootstrap récursif. Le wild bootstrap récursif, comme toute approche de bootstrap, nous permet d'avoir des résultats plus précis que la théorie asymptotique, pour des petits échantillons. Dans le souci de faciliter la lecture de notre rapport, nous avons consacré quelques pages à parler et illustrer le bootstrap ainsi que le wild bootstrap récursif.

Une fois les bases théoriques posées, nous avons décrit la démarche des tests effectués. La statistique de test utilisée est celle du ratio de variance. Deux approches ont été retenues. L'une est une approche non paramétrique pour tester l'hypothèse d'absence d'autocorrélation entre les log rendements (réels). Nous avons généré des pseudo log rendements, à partir de la série des log rendements observée, par le wild

bootstrap récursif. À chaque série artificielle de log rendements, nous avons calculé la statistique ratio de variance correspondante. Les pseudo statistiques ainsi obtenues nous ont permis de construire un intervalle de confiance bootstrap asymétrique. L'autre approche, quant à elle, utilise le modèle VAR pour dériver une équation du ratio de variance. Le modèle VAR a joué le rôle du processus générateur de données bootstrap. L'objectif reste à construire un intervalle de confiance bootstrap asymétrique plus fiable pour la statistique ratio de variance. Dans les deux approches, le test consistait à vérifier si les intervalles de confiance contenaient l'unité ou non. Si l'unité se trouve dans l'intervalle de confiance, la nulle de non prévisibilité ne pourra pas être rejetée. Les deux approches décrites ci haut nous ont donné des résultats favorables à la non prévisibilité des rendements financiers. Pour l'ensemble des horizons considérés ($k = 3, 6, 12, 24, 36, 48$ et 60 mois), les intervalles de confiance s'élargissaient avec l'horizon. Nos résultats sont proches de ceux de Malliaropoulos (1996), Politis, Romano et Wolf (1996) et Wolf (2000). Hodrick (1992), quant à lui, a affirmé trouvé de l'évidence sur la prévisibilité des rendements financiers en utilisant la statistique ratio de variance dérivée à partir d'un modèle VAR comme statistique de test. Notons qu'il pourrait être intéressant de considérer d'autres statistiques de test en conservant toujours l'approche de wild bootstrap récursif pour comparer les résultats ainsi trouvés. Une façon de choisir la bonne statistique de test est de prendre celle dont la puissance est la plus grande.

BIBLIOGRAPHIE

- Campbell, J., 1991: "A variance Decomposition for Stock returns", *Economic Journal*, 101, 157-179.
- Campbell, J, Lo A. W., et Mackinlay, A. C., 1997: "The Econometrics of Financial Markets", Princeton, NJ: Princeton University Press.
- Campbell, J. et Shiller, R. J., 1988: "Stock Prices, Earnings and Expected Dividends", *The Journal of Finance*, vol. 43, 661-676.
- Davidson, R., avril 1998: "Notes on the bootstrap", greqam.
- Efron, B., 1979: "Bootstrap Methods: Another look at the Jackknife", *Annals of Statistics*, vol.7, 1-26.
- Fama, E. et French, K., 1988: "Dividend Yields and expected Stocks Reurns", *Journal of Financial Economics*, 22. 3-25.
- Goetzmann, W. N., et Jorion, P., 1993: "Testing the Predictive Power of Dividend Yields", *Journal of Finance*, 48, 663-679.
- Gonçalves, S., et Kilian, L., 2002: "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form", working paper, Université de Montréal.
- Hodrick, R., 1992: "Dividend yields and expected stock returns: Alternative procedures for inference and measurement", *Review of Financial Studies*, vol.5, 357-86.

Kim, M., Nelson, C., et Startz, r., 1989: "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence", *Review of Economic Studies*, vol.58, 357-386.

Lettau, M., et Ludvigson, S., oct. 2001: "Measuring and Modelling Variation in the risk return trade off", Working paper, New York Univesity.

Lo, A. W., et Mackinlay, C., 1988: "Stock Market Prices do not Follow Random Walks. Evidence from a Simple specification Test", *The Review of Financial Studies*, vol.1, Issue1, 41-66.

Lütkepohl, H., 1991: "Introduction to Multiple Time Series Analysis", Springer-Verlag

Mackinnon, J.G. , 1998: "Bootstrap Testing in Econometrics", working paper, Queen's University at Kingston.

Malliaropulos, D., 1996: "Are long horizon stock returns predictable? A bootstrap analysis", *Journal of Business Finance and Accounting*, 23(1), 93-106.

Politis, D. N., 1998: "An Introduction to the Bootstrap, the Jackknife, and Other Household Items in a Statistician's Toolbox", *IEEE Signal Processing Magazine*, 39-54.

Politis, D. N., Romano, J. P., et Wolf, M., 1997 : "Subsampling for heteroskedastic time series", *Journal of Econometrics*, 81, 281-317.

Schwarz, G., 1978: "Estimating the Dimension of a Model", *Annals of Statistics*, vol.6, 461-464.

Wolf, M., 2000: "Stock Returns and dividend Yields Revisited: A New Way to look at an Old Problem", *Journal of Business and Economic statistics*, 18, 18-30.

ANNEXE I
Le critère de Schwarz

Soit la série chronologique multivariée de dimension k y_1, y_2, \dots, y_T , avec $y_t = (y_{1t}, \dots, y_{kt})'$, qui est généré à partir d'un processus $VAR(p)$

$$y_t = v + A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t \text{ tel que } \Sigma_u = E(u_t u_t')$$

En utilisant des arguments Bayesiens Schwarz (1978) dérive le critère (SC) ci-après pour choisir l'ordre p d'un modèle VAR.

$$SC(m) = \ln |\tilde{\Sigma}_u(m)| + \frac{\ln T}{T} \text{ (number of freely estimated parameters)}$$

$$SC(m) = \ln |\tilde{\Sigma}_u(m)| + \frac{\ln T}{T} mk^2$$

où $\tilde{\Sigma}_u(m)$ est l'estimateur du maximum de vraisemblance de Σ_u obtenu en imposant un modèle $VAR(m)$ aux données et T est la taille de l'échantillon.

L'ordre estimé $\hat{p}(SC)$ retenu est celui qui minimise le critère de sélection $SC(m)$.

Notons que SC est fortement convergent (Cf. Helmut Lütkepohl, Introduction to multiple Time Series Analysis, 1991, p.132).

ANNEXE II

Valeurs du critère de Schwarz (SC)

A : 1927 :02 à 1987 :11

Ordre du VAR	1	2	3	4
SC	-14.82972	-14.85146	-14.82577	-14.78017

B : 1952 :01 à 1987 :11

Ordre du VAR	1	2	3	4
SC	-16.89866	-16.82452	-16.73952	-16.67670

C : 1927 :02 à 1951 :12

Ordre du VAR	1	2	3	4
SC	-14.87255	-14.88913	-14.83699	-14.69855

D : 1927 :01 à 2001 :12

Ordre du VAR	1	2	3	4
SC	-15.10413	-15.13286	-15.11956	-15.07938

