TAMPERE EN YLIOPISTO
Taloustieteiden laitos

INVESTORS' BIASED EXPECTATIONS AND RETURNS PREDICTABILITY

−A TRADING RULE TEST

Kansantaloustiede
Lisensiaatintutkimus
Toukokuu 2009
Jukka Ilomäki

ABSTRACT

In this study a statistical technique is proposed which attempts to recognize statistical arbitrage opportunities. A certain kind of nonlinear cointegration approach is proposed, which we call temporary cointegration. In addition, unidirectional Granger causality is proposed to be an essential aspect for the forecastability of the technique. A mechanical trading rule based on the proposed technique produces economic profits in the testing period January 2000 − March 2007. In the trading rule test, the hit ratio for the forecasts becomes 80.8 % and it is statistically significant with p-value < 0.0001.

Because testing of the trading rule yields economically and statistically significant results, further research is needed to find possible causes for these economic profits. We examine this with two techniques. We determine (ex-post) the explanatory power of common risk factors that may produce the trading rule returns by multi risk factor model. In another approach, we forecast (ex-ante) with another procedure (macro-factor model and VAR without ECM) the same observations as in the trading rule and compare the results. After accounting for exchange rate effect, macroeconomic risk factors and lead-lag effect without ECM, the trading rule profits are still economically and statistically significant.

We suggest that when investors who invest to minor investing country assets receive ambiguous information concerning their assets, they may become confused in growing uncertainty. This may lead to a phenomenon where some investors take role models from large financial centers. In other words, there is a herding behavior with a lead-lag effect. The herding may start because investors are risk averters which act as a trigger for this particular herding behavior. If this herding behavior continues long and strong enough, this phenomenon may lead to a temporary relative dependency between particular assets. This may happen because investors suffer from an illusion of validity. There is a possibility that investors are overconfident about the success of their current investing behavior. Therefore, their actions affect the recency bias. We argue that these biases may lead to temporary cointegration with a lead-lag effect between particular asset prices. We propose that unidirectional temporary cointegration is consistent with the existence of herding with a lead-lag effect.

We demonstrate that in a two-asset economy where the investors have standard rational preferences but they may have biased beliefs, specific behavioral biases are sufficient conditions for existence of irrational herding with a lead-lag.

Keywords: temporary cointegration, unidirectional Granger causality, behavioral biases, herding.

TIIVISTELMÄ

Tutkimuksessa esitellään tilastollinen tekniikka, joka pyrkii tunnistamaan tilastollisia arbitraasimahdollisuuksia. Tietynlainen epälineaarinen yhteisintegroituvuus ehdotetaan, jota kutsutaan väliaikaiseksi yhteisintegroituvuudeksi. Ennustettavuuden kannalta ehdotetaan, että yksisuuntainen Granger-kausaalisuus on tärkeä aspekti. Kyseisen tekniikan perusteella muodostetaan mekaaninen sijoitussääntö, jonka avulla saadaan ylituottoja verrattuna normaalituottoon ajalla 1.1.2000–31.3.2007. Seuraavan askeleen suunnan ennustamisessa saavutetaan 80,8 % osumatarkkuus. Se on tilastollisesti merkittävä p-arvon ollessa < 0,0001.

Koska sijoitussääntö on tuottanut tilastollisesti ja taloudellisesti merkittäviä tuloksia, tehdään lisätutkimuksia ja pyritään löytämään mahdollisia selityksiä ylituotoille. Näitä tutkitaan kahdella tapaa. Määritetään jälkikäteen (ex-post) mahdollisten yleisten riskitekijöiden vaikutus sijoitussäännön tuottoihin. Lisäksi ennustetaan (ex-ante) samat sijoitussäännön määrittelemät kuukaudet sekä taloudellisten fundamenttien kehitystä kuvaavien tekijöiden avulla että ilman virheenkorjausmallia ja verrataan tuloksia. Näiden tekijöiden huomioimisen jälkeen sijoitussäännön tuotot ovat yhä taloudellisesti ja tilastollisesti merkittäviä.

Sijoitussäännön ylituottojen mahdolliseksi syyksi ehdotan behavioraalisia harhoja. Kun pieniin sijoitusmaihin sijoittavat sijoittajat saavat epäselvää informaatiota kyseisistä arvopapereista, he saattavat hämmentyä lisääntyvässä epävarmuudessa. Tällöin he voivat ottaa mallia suurista rahoituskeskuksista. Toisin sanoen voi syntyä johtaja-seuraaja laumakäyttäytymistä, joka ilmenee selvänä viiveenä kurssimuutoksissa. Laumakäyttäytyminen saattaa alkaa, koska sijoittajat ovat riskinkaihtajia. Tämä saattaa toimia laumakäyttäytymisen laukaisijana. Jos laumakäyttäytyminen jatkuu riittävän pitkään ja voimakkaana, tämä voi johtaa väliaikaiseen suhteellisten arvopaperihintojen keskinäiseen riippuvuuteen. Tämä saattaa tapahtua, koska sijoittajilla on validisuuden illuusio. He ovat mahdollisesti yli-itsevarmoja sen hetkisen sijoitusstrategian menestyksen suhteen. Näin heillä on viimeaikaisen tapahtuman ylipainottamisharha. Ehdotan, että nämä harhat voivat johtaa väliaikaiseen irrationaaliseen johtaja-seuraaja tyyppiseen yhteisintegroituvuuteen. Lisäksi ehdotan, että väliaikainen yksisuuntainen yhteisintegroituvuus on yhdenmukainen johtaja-seuraaja laumakäyttäytymisen kanssa.

Osoitan, että esitetyt behavioraaliset harhat ovat riittävä ehto irrationaaliseen johtaja-seuraaja tyyppiseen laumakäyttäytymiseen kahden assetin taloudessa, kun sijoittajat omaavat standardit rationaaliset preferenssit mutta harhaiset uskomukset.

Avainsanat: väliaikainen yhteisintegroituvuus, yksisuuntainen Granger-kausaalisuus, behavioraaliset harhat, laumakäyttäytyminen.

ACKNOWLEDGEMENTS

I would especially like to thank Professor Jari Vainiomäki, Docent Hannu Kahra, and Professor Hersh Shefrin for valuable guidance and comments. In addition, I would like to thank Professor Pentti Saikkonen, Professor Markku Lanne, Professor Tom Berglund, Professor Anders Löflund, Professor Hannu Vartiainen, Professor Avanidhar Subrahmanyam, Doctor Niklas Ahlgren, Professor Matti Keloharju, Doctor Markus Lahtinen, Professor Heikki Kauppi, Doctor Jan Antell and Licentiate Jani-Petri Laamanen for valuable comments. I am gratefully acknowledging financial support from Yrjö Jahnssonin säätiö, Säästöpankkiliiton tutkimussäätiö, Wallenbergin tutkimussäätiö and Tampereen yliopiston tukisäätiö.

CONTENTS

1. INTRODUCTION

A number of empirical studies have reported evidence on the predictability of stock returns. On an efficient financial market, investors are aware of any cross sectional or time variation in expected returns. Thus on an efficient financial market there are no arbitrage opportunities that would allow investors to achieve above average returns without accepting above average risk. This is our null hypothesis in the trading rule test developed later in this paper.

However, price changes are also produced by noise, which can be thought of as the opposite of information as proposed by Black (1986). He notes that noise trading is trading based on noise as if it was information, and it makes the financial markets imperfect. In various situations, this can produce correlated behavior on the financial markets. Behavioral finance studies how psychology affects financial decisions. It argues that financial market prices may deviate from fundamental values for long periods. According to behavioral finance studies, noise can be responsible for the sentiment based correlated behavior on the financial markets. If economic profits can be made based on these correlations, it is a contradiction to our null hypothesis. This is our alternative hypothesis in the trading rule test.

In this study, we propose a statistical technique that seems to recognize a temporary relative dependency between the assets concerned. A nonlinear cointegration approach with unidirectional Granger causality is introduced. If we can identify that this temporary cointegration with unidirectional Granger causality continues one-step-ahead, it may be an evidence of market inefficiency, because then one asset return can be used to forecast the other asset return.

Because testing of the trading rule yields economically and statistically significant results, we examine possible causes for these economic profits. We use two techniques to do this. The first one is to determine (ex-post) the explanatory power of common risk factors that may produce the trading rule returns by multi risk factor model as suggested by Lo (2008). We regress the trading rule returns on the following six factors: stocks, bonds, currencies, commodities, credit and volatility.

According to our regression results, these risk factors are not the source of economic profits.

Another technique is to forecast (ex-ante) with other procedure (which reflects possible reason for profits) the same observations as in the trading rule and compare the results. The trading rule test is done with US dollars. To find the exchange rate effect for (the identified lead-lag months) the trading rule profits returns are also calculated in local currencies. The second possible explanation for economic profits is that they are caused by macroeconomic variables such as developments of dividend yield, local short-term interest rate and local long-term bond yield. To find the forecast power of these variables for asset returns we construct a factor model proposed by Solnik (1993). Thus, we use developments of dividend yield, local short-term rate and local bond yield as proxies for country specific fundamentals. The third possibility for the trading rule economic profits is based on the behavioral finance literature. Lo and McKinlay (1990) argue that there may be some absolute lead-lag effect between the assets concerned. That is the cross autocorrelation puzzle. To clarify this phenomenon the same identified lead-lag months as in the trading rule are tested without ECM (error correction mechanism) in the model.

We suggest an explanation for the remaining economic profits, which is based on the behavioral finance literature. We assume that investors have standard rational preferences but they may have biased beliefs. We suggest that some biased expectations may lead to a phenomenon where some investors take role models from large financial centers, which lead to a temporary lead-lag effect between relative asset prices. In other words, we propose the following: If investors who invest to minor investing country assets receive ambiguous information, which leads to growing information uncertainty about future returns on their assets they may take temporary role models from large financial centers for their future investing decisions. Risk aversion may act as a trigger. With the recency bias, the illusion of validity and overconfidence this may lead to temporary relative dependency between particular assets. When we identify this, it is an evidence of irrational herding with a lead-lag. Thus, we propose that temporary cointegration with unidirectional Granger causality is a test of existence of (rational and irrational) herding with a lead-lag effect. In a sense, irrational herding can be also called financial contagion. For example, Cipriani & Guarino (2008) define financial contagion as a phenomenon where there is correlation between two assets that exceeds the correlation between the fundamentals of the assets. However, in the study we prefer the term irrational herding.

The proposed technique suggests that the link between temporary cointegration with unidirectional Granger causality and the forecastability is the following. When we identify the forecastability with the technique proposed, it means that this technique suggests that the temporary cointegration with lead-lag effect continues to the next step ahead. Thus, the probability of continuity of unidirectional temporary cointegration to the next step is greater than 0.5. This technique therefore makes a forecast about the likely direction of one-step-ahead price change of the follower. The lead-lag effect suggests that the follower is forecastable, because the relative dependency forces the follower to follow the leader's relative price and so the equilibrium continues on the next step ahead.

In addition, we demonstrate that in the two-asset economy where the investors have standard rational preferences but they may have biased beliefs, specific behavioral biases are sufficient conditions for existence of irrational herding with a lead-lag. Therefore, we suggest that assets prices may start to deviate from their fundamental values or from their different kinds of sentiment pricing in our case because of those investors' specific behavioral biases in times of growing uncertainty. These biases may create irrational herding behavior and may lead to temporary cointegration with unidirectional Granger causality. Thus, in that period specific sentiment weighs more than the economic fundamentals in investors' decisions. This situation persists until debiasing starts to affect and the fundamentals or different kinds of sentiment again start to outweigh this specific sentiment. In addition, global fundamentals should create simultaneous comovements between assets, not with a lead-lag effect. However, if country specific fundamentals create herding behavior with lead-lag effect among assets, this phenomenon should be identifiable with factor models (which take account of those fundamentals).

The testing data is MSCI (Morgan Stanley Capital International) country indices as monthly data. The calibration period is January 1990 – December 1999 and testing period is January 2000 – March 2007. All indices have been noted in US dollars.

The structure of this study is as follows. In the next section, we present a literature review of theoretical and empirical studies of return predictability. In Section 3, we propose temporary cointegration with unidirectional Granger causality and its connection to investors' behavior. In addition, in this section we present irrational herding with a lead-lag model in a two-asset economy. Testing of the trading rule and the results are presented in Section 4. In Section 5 we discuss possible causes for the trading rule economic profits and Section 6 concludes.

2 LITERATURE REVIEW OF RETURN PREDICTABILITY

2.1 MARKET EFFICIENCY TODAY

Malkiel (2003) argues that financial markets are assumed to be information efficient. This means that no arbitrage opportunities exist that would allow investors to achieve above average returns without accepting above average risk.

Let us define asset return as

$$r_t = (P_t + D_t - P_{t-1}) / P_{t-1} \qquad (2.1)$$

where P is asset price, r is return and D is dividend. Harrison and Kreps (1979) demonstrate that there cannot be riskless or pure arbitrage opportunity (PAO) in a stable economy. A PAO is possible if a zero cost trading strategy is feasible that offers the positive expected payoff or economic profits with no possibility of loss. Harrison and Kreps show that only in an arbitrage free economy stochastic discount factor exist. The discount factor Q_{t+1} is stochastic because it is not known with certainty at time t. This means that

$$P_t = E[Q_{t+1}(P_{t+1} + D_{t+1})|\theta_t], \qquad (2.2)$$

where $E[. | \theta_t]$ is conditional expected value operator, P is price of asset, D is dividend and Q_{t+1} is next period discount factor. This way stochastic discount factor always affects asset pricing, but because it is unobservable equation (2.2) translates into simple martingale condition

$$P_t = E[(P_{t+1} + D_{t+1})|\theta_t].$$

Thus it implies that the best forecast of tomorrow's price plus dividend is today's price. Subtracting a risk free return (for example three months T-bill) from Equation (2.1) we get risk adjusted return R_t. Now we can define the risk adjusted return for the next period as

$$R^*_{t+1} = R_{t+1}Q_{t+1},$$

where Q_{t+1} is a stochastic discount factor. Granger & Timmermann (2004) note that informational efficiency implies that with respect to information set θ_t, the following equation must hold:

$$E(R^*_{t+1}| \theta_t) = 0.$$

This means that the next period (t+1) risk adjusted return with respect to information set θ_t cannot be predicted, because the stochastic discount factor Q_{t+1} is unobservable.

Market efficiency means also that there are no statistical arbitrage opportunities in the financial markets. To see this we review the following articles from the field.

Cochrane and Saá-Requejo (2000) propose that an efficient market also rules out investment opportunities with continuously high Sharpe ratios. Bernando and Ledoit (2000) propose ruling out investment opportunities that offer high gain-loss ratios when compared to risk adjusted return.

Bondarenko (2003) proposes that the developed economy also rule out statistical arbitrage opportunities (SAO). In a finite horizon economy, there exists SAO if a zero cost trading strategy can be made for which the expected payoff is positive and the conditional expected payoff in each final state of the economy is nonnegative. Unlike a PAO, a SAO may have negative payoffs provided that the average payoff in each final state is nonnegative. More formally, Bondarenko defines a statistical arbitrage as a zero cost trading strategy with a payoff

$$Z_T = Z(I_T)$$

is a statistical arbitrage if

$$E[Z_T \mid I_0] > 0$$

and

$$E[Z_T \mid I_0^{eT}] \geq 0$$

for all e_T, where $I_0^{eT} := (I_t \,;\, e_T) = (e_1, e_2, \ldots e_t \,;\, e_T)$, Z is value of payoff at time t, I is information set and e_t is a random variable. Bondarenko defines that there is a finite number of trading dates, indexed by t = 0,1,….T. At time t, the state of the economy is represented by a random variable e_t. The history of states up to time t determines the market information set $I_t = (e_1, e_2, \ldots e_t)$. Bondarenko notes that the stochastic discount factor exists only if there are no statistical arbitrage opportunities in an economy. Testing this restriction can be viewed as model free, because it requires no parametric assumptions about the true equilibrium model. It continues to hold when investors' beliefs are mistaken and it can be tested in samples affected by selection biases such as the peso problem. Bondarenko shows that this resolves the joint hypothesis problem when market efficiency is tested.

2.2 LITERATURE REVIEW OF TESTING RETURN PREDICTABILITY

Many empirical studies have reported evidence on the predictability of stock returns, which may lead to market inefficiency. The main reported phenomena that may lead to inefficiency are investors' overreaction and underreaction to unexpected information. The study in experimental psychology by Kahneman & Tversky (1974) found that people tend to overreact to unexpected and dramatic events. De Bondt &

Thaler (1985) were the first ones to try formally test this cognitive misperceptions theory in the financial markets. They argue that investors overreact to both bad news and good news. Therefore, overreaction leads past losers to become undervalued and past winners to become overvalued. Further studies of overreaction have been provided by Summers (1986), Chan (1988), Fama & French (1988a), Lehmann (1990), Jegadeesh (1990), Lo & MacKinlay (1990), Jegadeesh & Titman (1995) and Subrahmanyam (2005) among others.

Conservatism bias was introduced by Edwards (1968). He suggests that people underestimate new information in updating their priors. Lo & MacKinlay (1988) found a positive serial correlation in weekly returns which is statistically and economically significant. If investors act in this way, stock prices tend to slowly adjust to new information as proposed by Jegadeesh & Titman (1993). This phenomenon can be translated to momentum profits, which arise because investors underreact to ranking period information, which is gradually incorporated into stock prices during the holding period. Further studies have provided by Fama & French (1988b), Conrad & Kaul (1998), Jegadeesh & Titman (2001) and Zhang (2006) among others.

Factor models detect phenomena where economic fundamentals cause return predictability. Fama & French (1992 & 1993) note that firm's size and book to market ratio predict stock returns. They argue that these variables can explain the differences in average returns across stocks because these variables characterize the riskiness of the firm. This means that return predictability in this case is normal compensation for the risk and so firm's size (negatively correlated with stock returns) and book to market ratio (positively correlated with stock returns) are proxies for fundamentals. Fama & French (1988b), Campbell (1991) and Cochrane (1992) argue that dividend yields vary over time, dividend yield variability helps to predict the risk premium of stock prices. They note that aggregate dividend yields strongly predict excess returns on a longer horizon. Solnik (1993) argues that macroeconomic variables such as developments of dividend yield, local short-term interest rate and local long-term bond yield predict stock returns

Technical analysts attempt to forecast prices by the study of past prices. Technical analysts believe that shifts in supply of and demand for stocks can be detected in charts of market actions. A sophisticated technical idea is to use cointegration analysis to forecast price changes. In this study we use temporary cointegration approach with unidirectional Granger causality.

2.2.1 COINTEGRATION ANALYSIS RELATED STRATEGIES

Granger (1986) argues that if a pair of stock prices were cointegrated, it would be a clear evidence of market inefficiency. Gatev, Goetzmann and Rouwenhorst (2006) tested an investment strategy called “pairs trading”, which is closely related to cointegration based trading rules. Pairs trading means basically finding two stocks whose prices have moved together historically. When the spread between these stocks widens the strategy suggests buying (selling) the loser (winner). They chose from the daily data the pairs of stocks that have moved together for last 12 months. They opened a position in a pair when prices diverged by more than two historical standard deviations, as estimated during the pairs’ calibration period. They closed their position after six months. They found an average 11 % annual economic profits in USA stocks between 1962 and 2002. However, they note that these economic profits have been declining since beginning of the nineties.

Kasa (1992) observed that USA, Germany, Japan, United Kingdom and Canada stock indices share one common stochastic trend. Thus according to Kasa they are cointegrated. However, Richards (1995) criticized Kasa’s results because his use of asymptotic critical values in the cointegration tests is not appropriate. However, Richards finds predictability between national stock returns, but he does not find significant evidence of cointegration.

Alexander (1999) developed a trading rule technique based on cointegration analysis. This technique detects stocks that are cointegrated with the reference index and recommends concentrate to invest only on those stocks. Alexander called this technique cointegration based index tracking.

From cointegration based index tracking Alexander and Dimitriu (2005a) developed a trading strategy which has contributed economic profits. The main advantage of their idea is the reduction of the noise present in stock returns. According to Alexander and Dimitriu (2005a) a cointegration based index tracking portfolio structure is more stable than a benchmark index. They argue that especially in volatile markets, when returns are low and prices are moving towards new levels, the strategy produces consistent economic profits. The basic idea of the trading strategy is to speculate that the error correction mechanism will bring stock prices back to long-term equilibrium. According to the market neutral strategy (Alexander & Dimitriu 2005b) with stocks that have been revealed to be cointegrated with benchmark index it is advisable to open a position which depends on location with respect to long-term equilibrium. When the selected stock is above (below) the long-run equilibrium it should be sold (bought).

Dunis and Ho (2005) tested a trading strategy similar to Alexander and Dimitriu (2005b) with a small recent EuroStoxx50 sample (January 1999 − June 2003). They conclude that cointegration based portfolios add to an investor's value. Our technique falls into the same category as Dunis & Ho and Alexander & Dimitriu.

3. TEMPORARY COINTEGRATION

3.1 COINTEGRATION

According to Engle and Granger (1987) non-stationary I(1) time series $\{Y_t\}$ and $\{X_t\}$ are cointegrated if there exists a stationary linear combination $\beta' x_t$, where β' is the cointegration vector. They argue that the components of the vector x_t are cointegrated CI(1,1) if the vector β' exists such that

$$Z_t = \beta' x_t \sim I(0),$$

where $\beta' \neq 0$ and I(0) means stationary. (Engle & Granger 1987)

If x_t has **n** components, there may be more than one cointegration vector β'. Then there are **r** linearly independent cointegrating vectors, with $\mathbf{r \leq n\text{-}1}$, which are gathered together into the **n x r** array β'. By construction the rank of β' will be **r,** which will be called the cointegrating rank. (Engle & Granger 1987)

According to the Granger representation theorem, when a stationary linear combination can be found, there must also be an error correction mechanism. This can be illustrated such that

$$D(B)\,(1 - B)x_t = -\alpha Z_{t-1} + e_t,$$

where $\alpha \neq 0$, $Z_t = \beta' x_t$, $D(0) = I$, $D(1)$ has all elements finite, B is the lag operator and e_t is a stationary multivariate residual vector. In time series $\{X_t\}$ and $\{Y_t\}$ the error correction mechanism can be shown as a VAR form such that

$$\Delta X_t = -\alpha_x(X_{t-1} - \beta Y_{t-1}) + \text{lagged } \Delta X_t\,,\ \Delta Y_t + e_{xt} \qquad (3.1)$$

and

$$\Delta Y_t = \alpha_y(X_{t-1} - \beta Y_{t-1}) + \text{lagged } \Delta X_t\,,\ \Delta Y_t + e_{yt}. \qquad (3.2)$$

In Equation (3.1) α_x is the response power of the series $\{X_t\}$ to long-term equilibrium, and α_y is the response power of (Equation 3.2) $\{Y_t\}$. Because in our trading rule it is required that only series $\{X_t\}$ has a statistical significant reaction power to the equilibrium errors, then α_y must be ≈ 0. It follows that series $\{Y_t\}$ does not have a statistical significant reaction power to the equilibrium space $(X_{t-1} - \beta Y_{t-1})$, when Equation (3.2) can be written as

$$\Delta Y_t = e_{yt}. \qquad (3.3)$$

Thus, {Yt} is weakly exogenous in the cointegration system. To see G-causality through ECM (Granger 1988), consider the model

$$X_t = AQ_t + X_{1t},$$

and

$Y_t = Q_t + Y_{1t}$,

where $Q_t \sim I(1)$, $\{X_t\}$ and $\{Y_t\}$ both are integrated of the order 1, but cointegrated, because

$Z_t = X_t - AY_t$,

where $\beta = (1 - A)'$, $x_t = (X_t, Y_t)'$ and X_{1t} & Y_{1t} are both white noises. The error correction model is then (equations 3.1 and 3.2). Let Q_t be the random walk such that $\Delta Q_t = a_t$. Hence a_t, X_{1t} and Y_{1t} are all white noises. Then,

$Z_t = X_t - AY_t = X_{1t} - AY_{1t.}$

Now consider

$\Delta X_t = -X_{1t-1} + Aa_t + X_{1t}$,

where $-X_{1t-1}$ is the forecastable part of ΔX_t and Aa_t and X_{1t} constitute the one-step forecast error of forecast made at time t-1. Granger continues that the forecast $-X_{1t-1}$ is not directly observable but is correlated to Z_{t-1}, which results in the G-causality.

3.2 UNIDIRECTIONAL GRANGER CAUSALITY

In this study, we propose that unidirectional (one-way) Granger causality (1969), offers together with the Granger representation theorem and with temporary cointegration, a key to statistically significant predictions. Granger (1980) defines Granger causality (G-causality) as a state that satisfies two conditions: The cause preceded the effect and a causal series had information about the effect that was not contained in any other series according to the conditional distributions. Granger argues that the implication of these statements is that using the cause produces a superior forecast of the effect. In other words, **the basic definition of G-causality is a statement about predictability.** More formally, Granger (1969) means by causality that a series $\{Y_t\}$ has a causal effect on a series $\{X_t\}$, when

$\sigma^2(X_t|\mathbf{X}) > \sigma^2(X_t|\mathbf{X},\mathbf{Y})$,

where

$\mathbf{Y} = \{Y_{t-1}, Y_{t-2},\ldots, Y_{t-r}\}$

and

$\mathbf{X} = \{X_{t-1}, X_{t-2},\ldots, X_{t-r}\}$,

where σ^2 is the forecast error. Granger (2003) points out that this does not mean that exogeneity is the same as G-causality but rather Y can be called as weakly exogenous. Weakly exogenous in the cointegrated systems means that Y (in our case) is weakly exogenous in the long-run. Thus, it means that Y does not respond to equilibrium error. Granger (1988) argues that **in the error correction model, there are two possible sources of G-causality of X_t by Y_{t-1}, either through the error correction mechanism (ECM) or through ΔY_{t-1}.**

Granger causality in (two variables) cointegration relationships works usually back and forth. The reason for this is that usually they both react via the error correction mechanism to the deviations from the long run equilibrium. In this study, we are interest in special cases where unidirectional Granger causality is affecting among cointegrated variables. That is, one cointegrated variable is an effect and the other is a cause. Granger (1986) argues that as long as the two variables are cointegrated, causality must exist in at least one direction, as one variable can help forecast the other.

3.3 REVIEW OF NONLINEAR COINTEGRATION

Extensions of standard cointegration have been growing rapidly. If the data are cointegrated, error correction models are estimated. However, many economic relationships are often nonlinear; this phenomenon leads to a situation where linear cointegration is not observed.

Balke and Fomby (1997) propose a model in which the cointegration relationship between variables turns on and off. They illustrate this on and off behavior as a threshold cointegration in which the series are cointegrated if they move too far away from the equilibrium relationship but are not cointegrated as long as they are relatively close to the equilibrium. Further developments of the threshold cointegration have been accomplished for example by Tsay (1998), Lo & Zivot (1999),Hansen & Seo (2002) and Saikkonen (2008).

Another kind of nonlinear cointegration is proposed by Granger & Yoon (2002). When the components of data series are cointegrated, they argue that the data series have hidden cointegration. Granger & Yoon define hidden cointegration as follows: Consider two random walks

$$X_t = X_{t-1} + \varepsilon_t = X_0 + \sum\varepsilon^+_i + \sum\varepsilon^-_i$$

and

$$Y_t = Y_{t-1} + \eta_t = Y_0 + \sum\eta^+_i + \sum\eta^-_i,$$

where X_0 and Y_0 denote initial values, ε_t and η_t are white noises, $\sum\varepsilon^+_i$ and $\sum\eta^+_i$ are sums of positive shocks, $\sum\varepsilon^-_i$ and $\sum\eta^-_i$ are sums of negative shocks. For example, when the sums of positive shocks are cointegrated, both X and Y are subject to common positive shocks. On the other hand, in this case the sum of negative terms is not cointegrated. Granger & Yoon argue that this hidden cointegration will not be utilized if we are interested only in the cointegration between X and Y.

The third kind of nonlinear cointegration is proposed by Granger and Siklos (1997). They propose the concept of regime sensitive cointegration. They argue that it is realistic to assume that some series are cointegrated in some periods and not in others. They suggest that some economic relationships are best thought of as occasionally falling in or out of equilibrium because of some major policy shift or event. Granger and Siklos define regime sensitive cointegration as follows: Let us examine the long run relationship between time series $\{X_t\}$ and $\{Y_t\}$ such that both are integrated of the order 1, denoted I(1). Suppose that cointegration exists only during certain periods; otherwise, the series are not cointegrated. To demonstrate this, consider the following model:

$X_t = AQ_t + \lambda_{1t}W_t + \alpha_1 Z_t$

and

$Y_t = Q_t + \lambda_{2t}W_t + \alpha_2 Z_t$,

where W_t and Q_t are ~ I(1) and not cointegrated, but Z_t ~ I(0), $\alpha_i > 0$, with $\alpha_1 - A\alpha_2 = 1$. If $\lambda_{1t} = \lambda_{2t} = 0$ or $\lambda_{1t} - A\lambda_{2t} = 0$,, then X_t and Y_t will be cointegrated. But this property will **not** hold in general, because there are two stochastic trends (W_t and Q_t). Thus, the parametric (λ_{1t} & λ_{2t}) are assumed to vary over time.

3.4 REVIEW OF THE ROLLING WINDOW METHOD

In the simplest form of the temporary cointegration, we propose that the rolling window technique reacts to the breaks, which cause variables to fall in or out of cointegration. The rolling window goes back to the early statistical quality control literature of Shewhart (1939). An example of its more recent use in the finance literature, Shiller and Fair (1990) argue that using rolling estimation forecasts is important because in so doing we are producing better forecasts as time progresses. Lo (2008) notes that the rolling window method can address the ubiquitous issue of nonstationarity that affects most financial studies; time-varying means, volatilities, and general market conditions can be captures to some degree by using rolling windows. Thus, the rolling window is a way to prepare for instability of the data. Granger & Siklos (1997) note that in regime sensitive cointegration it is important to identify structural breaks that cause regime shifts.

Heracleous, Koutris and Spanos (2005) define the rolling window method such that

$\{R_t\}_{\ t = 1,2,\ldots\ldots\ldots..n}$

is a random process, θ be the unknown parameter to be estimated and $\hat{\theta}$ = g(R_t) is an estimator based on the process. In addition, let

$P_R = \{P_{R_i}\}_{i \in I}$ be a partition of the process,

such that

$P_{R_{ti}} = \{R_t : t \in [ti, ti-1 + l\,]\}$

for ti = 1,2,……..n - (l -1)

and

l = fixed window size. Then the rolling estimator $\hat{\theta}_{rti}$ of the unknown parameter θ is defined as

$\hat{\theta}_{rti} = g(R_{ti})$

for ti = 1,2,……..n - (l -1).

This provides us with a sequence of estimates instead of just one estimate for $\hat{\theta}$. Each estimate is based on the same number of observations, because both the starting date and the ending date of the window move by one period, when the window moves ahead.

Swanson (1998) was among the first to apply the rolling window method to cointegration analysis. Swanson studied money (M1) and output long run linkage with the rolling window. He used this method because he allows that the system may be evolving over time, thus avoiding the problem of estimating potentially unstable cointegration relations. In his study, the degree of cointegration varies across samples, but variables remain cointegrated all the time in the data. Thus rolling cointegration is a way to study the stability of cointegration.

Rangvid and Sorensen (2002) propose using time varying multivariate cointegration techniques to analyze the dynamics of convergence of variables in process of converging. They use rolling cointegration tests to find out possible declining numbers of common stochastic trends. They argue that when the rolling estimation is used, the power of the tests remains constant. This is an advantage when rolling window method is used and makes it possible to analyze how the process evolves over time. Rangvid and Sorensen studied how the convergence of five ERM exchange rates developed during the ERM period.

Brada, Kutan and Zhou (2005) propose that the technique of rolling cointegration detects time varying estimates of the convergence of macroeconomic variables within EU and between transition economies and the EU. They argue that using rolling cointegration tests explicitly takes into account the possibility that data series are more cointegrated during some parts of the sample period but less so or not at all during other parts**.** Additionally, Brada et al. argue that they deal with the possibility of gradually time varying cointegration by using rolling cointegration, a technique that explicitly allows for changes in the relationship between systems of variables.

According to Engle and Granger (1987) cointegration is a long run equilibrium relation. Thus, short-term analyses, which rely on cointegration, may therefore have inconsistency with the theory, because the reaction of the error correction mechanism may be slow. When in short-term studies we concentrate only on predicting the direction of change, the situation is different. Swanson and Zeng (1998) propose that the rolling cointegration based forecast has predictive power in one-step ahead predictions. They investigated the forecast ability of future spot prices using the term structure of future prices of four commodities (S&P 500 index, treasure bond, gold and crude oil). It can be assumed that all of their data has stable cointegration. They find that the error correction mechanism works best when only the direction of changes is forecasted.

We use the rolling window method to detect structural breaks causing situations where $(\lambda_{1t} - A\lambda_{2t}) = 0$**.** That is, integrated variables X_t and Y_t are temporarily cointegrated.

3.5 DEFINITION OF UNIDIRECTIONAL G-CAUSALITY TEMPORARY COINTEGRATION AND ONE-STEP-AHEAD FORECASTING

For the basic definition of temporary cointegration, we follow Granger & Siklos (1997). Consider the following model:

$$X_t = AQ_t + \lambda_{1t}W_t + \alpha_1 Z_{xt} \tag{3.4}$$

and

$$Y_t = Q_t + \lambda_{2t}W_t + \alpha_2 Z_{yt}, \tag{3.5}$$

where Q_t and W_t are I(1) and not cointegrated, but Z_{it} is I(0), $\alpha_i > 0$, with $\alpha_1 - A\alpha_2 = 1$. If $\lambda_{1t} - A\lambda_{2t} = 0$, then X_t and Y_t will be cointegrated. However, this property will not hold in general, because there are two stochastic trends (Q_t and W_t).

To see temporary cointegration with unidirectional Granger causality we multiply equation (3.5) by A and subtract from (3.4) to get

$$Z_t = X_t - AY_t = AQ_t - AQ_t + \lambda_{1t}W_t - A\lambda_{2t}W_t + \alpha_1 Z_{xt} - A\alpha_2 Z_{yt.}$$

From this we get

$$Z_t = X_t - AY_t = (\lambda_{1t} - A\lambda_{2t})W_t + \alpha_1 Z_{xt} - A\alpha_2 Z_{yt,}$$

where Q_t stands for one (the same in both series X_t and Y_t) stochastic trend. $Z_t \sim I(0)$. If $(\lambda_{1t} - A\lambda_{2t}) = 0$, then X_t and Y_t are temporarily cointegrated.

Let us examine cases when ΔX_{t+1} is cointegrated and when it is not. Let us define that Q_t and W_t are pure random walks. Thus, $\Delta Q_t = a_t$, $\Delta W_t = b_t$ and a_t & b_t are white noises. Now consider

$$\Delta X_{t+1} = A\Delta Q_{t+1} + \lambda_{1t}\Delta W_{t+1} + \alpha_1 \Delta Z_{xt+1.}$$

From this we get

$$\Delta X_{t+1} = Aa_{t+1} + \lambda_{1t+1}b_{t+1} + \alpha_1 Z_{xt+1} - \alpha_1 Z_{xt}$$

$-\alpha_1 Z_{xt}$ is the forecastable part of ΔX_{t+1} and $Aa_{t+1,}$ $\lambda_{1t+1}b_{t+1}$ and $\alpha_1 Z_{xt+1}$ constitute the one-step forecast error of forecast made at time t. The forecast $-\alpha_1 Z_{xt}$ is not directly observable but is correlated to Z_t, which results in the G-causality. Thus, $-\alpha_1 Z_{xt}$ can be forecasted with equilibrium error Z_t. That is,

$$Z_t = \alpha_1 Z_{xt} - A\alpha_2 Z_{yt}$$

if X_t and Y_t are temporarily cointegrated and

$$Z_t = (\lambda_{1t} - A\lambda_{2t})W_t + \alpha_1 Z_{xt} - A\alpha_2 Z_{yt}$$

otherwise. Thus, if X_t and Y_t are temporarily cointegrated

$$Z_{xt} = 1/\alpha\ [Z_t + A\alpha_2 Z_{yt}], \tag{3.6}$$

and otherwise

$$Z_{xt} = 1/\alpha\ [Z_t + A\alpha_2 Z_{yt} - (\lambda_{1t} - A\lambda_{2t})W_t\ ]. \tag{3.7}$$

It can be easily seen from the equations (3.6) and (3.7) that the variance of the ΔX_{t+1} is larger when temporary cointegration is off. Because W_t is a random walk, its variance is growing linearly in time. It implies that when temporary cointegration exits the forecast error of ΔX_{t+1} is smaller.

We use the rolling window method to detect structural breaks causing situations where $(\lambda_{1t} - A\lambda_{2t}) = 0$. That is, integrated variables X_t and Y_t are temporarily cointegrated. Let us define

$\{\varphi_t\}$,

where $\{\varphi_t\} = (\varphi_1, \varphi_2, \varphi_3, \ldots\ldots\ldots\ \varphi_t)$. This constructs from series $\{Y_t\}$ and $\{X_t\}$ and let (Y_t, X_t) be relative dependency (that is possible cointegration), which we are modeling. Then

$\varphi_{\tau i}$

is the changing part entity such that

$\varphi_{\tau i} = \{\varphi_t : t \in [\tau i, \tau i\text{-}1 + l\,]\}$

where $\tau i = 1,2,\ldots\ldots..n - (l\,\text{-}1)$

and l is the fixed window size.

The main goal for the temporary cointegration is to find when time series $\{X_t\}$ is forecastable. A difference between our approach and Brada et. al. (2005) is that they use the rolling cointegration approach to estimate the stability of cointegration rank. They do not use it to forecast anything.

To generalize temporary cointegration we can define that window size l is not constant over time. Fixed window size is rather a primitive approach to optimize forecastability, because it can be assumed that structural change is not constant over time. Additionally, according to Pesaran and Timmermann (2002) immediately after a break the window will tend to be too long and further away from the break the window will be too short. The problem is that no further information is used to determine possible time variation in window size.

To forecast temporary cointegration with the rolling window let us define that

$\varphi_{\tau i}$ = window i

and

$\varphi_{\tau i+l}$ = window i + one step ahead.

We define that $\varphi_{\tau i+l}$ is the same sample as $\varphi_{\tau i}$ except that it also includes an unknown observation $\tau i+l$. Now let us assume that we find $\varphi_{\tau i}$ (where $t \in \varphi_{\tau i}$) to be cointegrated in the following way

$$\Delta X_{t+1} = \mu - \alpha_x (X_t - \beta_i Y_t) + \lambda_1 \Delta X_t + \kappa_1 \Delta Y_t + e_{xt+1} \qquad (3.8)$$

and

$$\Delta Y_{t+1} = \mu + \alpha_y (X_t - \beta Y_t) + \lambda_2 \Delta X_t + \kappa_2 \Delta Y_t + e_{yt+1}. \qquad (3.8')$$

Now consider $f(\varphi_{\tau i+l})$ is a one-step predictive distribution and ΔX_{t+1} actually occurs, then forecast error will itself be a distribution

$$f[\Delta X_{t+1} - (\mu - \alpha_x (X_t - \beta_i Y_t) + \lambda_1 \Delta X_t + \kappa_1 \Delta Y_t)] = f(e_{xt+1}), \qquad (3.9)$$

where f means forecast available at time t.

We should note that in the simple form of temporary cointegration (rolling window with a constant *l*) successive samples are the same except for the first and last observation. Therefore this overlapping quality has to be taken into account in forecasting. For example, in this study the main goal is to forecast the direction of one-step price changes. Also the cointegrated relation additionally defined by Engle and Granger (1987) is a long run equilibrium. When cointegration analysis is used to forecast one-step ahead, there is the possibility that the error correction mechanism has no time to affect. However, according, for example, Swanson & Zeng (1998) and Alexander & Dimitriu (2005a) the cointegration relation has forecasting power even in one-step-ahead predictions. Swanson and Zeng find that this is best seen when the goal is to forecast only the direction of changes.

We argue that we can forecast the direction of changes by studying the developments of the modeling criteria (see Subsection 3.6). We assume that the more statistically significantly modeling criteria test results are, the greater is the probability that temporary cointegration with unidirectional Granger causality will continue to the next-step-ahead and we can compute predictive distribution $f(\varphi_{\tau i+1})$. We have a belief that the probability of continuity is > 0.5 when the modeling criteria boundaries are reached.

In other words, when we identify temporary cointegration with unidirectional Granger causality and construct equations (3.8 & 3.8'), we can calculate with maximum likelihood method the predictive distribution $f(\varphi_{\tau i+1})$ to the next step. The more significant the modeling criteria test results are, the stronger belief we have about continuity of the unidirectional cointegration to the next-step-ahead and thus the correctness of predictive distribution $f(\varphi_{\tau i+1})$ (equation 3.9). We use loss function to analyze our forecasting results and to update our beliefs.

Summary 1. When we identify forecastability with the proposed technique, it means that this technique suggests that the temporary cointegration with lead-lag effect continues to the next step ahead. Thus, the probability of continuity of unidirectional temporary cointegration to the next step is > 0.5. The technique makes a forecast about the direction of one-step-ahead price change of the follower. The lead-lag effect suggests that the follower is forecastable, because the relative dependency forces the follower to follow the leader's relative price in order for the equilibrium to continue to the next step ahead. This makes the direction of follower's one-step-ahead price change forecastable.

3.6 ACCEPTABLE MODELING CRITERIA FOR THE TEMPORARY COINTEGRATION TO THE TRADING RULE

In this study the calibration period is January 1990 – December 1999 for defining fixed window size. (It means that in this period we have chosen the fixed window size and following other criteria for the out of sample period January 2000 – March 2007.) Two and a half years (or 30 months) is chosen as the window size.

1) The testing of cointegration from the samples is done by Johansen's test (Johansen, 1988 & 1991) **where the critical p-value for trace test is 0.070**. (for more information see appendix I)

2) Granger causality test is done by likelihood ratio test for **long-run weak exogeneity where the required p-value for the restriction that (H_0) $\alpha_y = 0$ is > 0.15 and for the restriction that (H_0) $\alpha_x = 0$ p-value is < 0.0725.** In other words, when the LR-test gives a p-value > 0.15 for the restriction that $\alpha_y = 0$ the null hypothesis cannot be rejected and we can conclude that sub series $\{X_t\}$ does not Granger cause sub series $\{Y_t\}$. At the same time we require that the p-value for the restriction that $\alpha_x = 0$ is < 0.0725, so that the null hypothesis can be rejected, which implies that sub series $\{Y_t\}$ may Granger-cause $\{X_t\}$. This test is included in the Johansen methodology (for more information see appendix I and Johansen 1988 & 1991 and Johansen & Juselius 1990).

We use full information maximum likelihood method to estimate equations (3.8) and (3.8') to get coefficients α_x, α_y, β_i, λ, κ and constant μ. We construct both equations because we consider temporary cointegration relationships as very delicate system where on and off quality affects to the relationship of the variables.

3) In the final model the p-value of t-statistic for ECM variable (α_x) in series $\{X_t\}$ is required < 0.0725. Thus, we confirm the correctness of the results from the previous test (G-causality). The choice of the final forecast model is done by AIC. This means that we choose lags of variables with Akaike Information Criteria, where model that produces the smallest AIC-value is chosen.

4) Because a forecastable model is required to filter out white noise, **diagnostic tests are needed, where all p-values are ≥ 0.05.** The autocorrelation test is done by (Breusch-Godfrey) LM-test. ARCH character is tested with LM-test with squared residuals. Heteroskedasticity is tested with White's test. The normality distribution of the residual is tested with a small sample adjusted Jarque & Bera test.

If any of the (4) tests cannot reach boundaries (required p-values) it means that we do not use that rolling window sample in the trading rule test.

In other words, we do the following: In the calibration period (January 1990 – December 1999), we define sample size and boundaries for p-values of modeling criteria by using lost function. That is, we examine trade off between the hit ratio (see Chapter 4.3.1) and number of forecasts. Our goal is to get the p-value of PT-test as low as possible during the calibration period. Thus, we study the hit ratio and modeling criteria boundaries with different window sizes (from 20 to 80 observations). We chose 30 observations for the fixed window size. Then, we start real time testing. Because we chose 30 months as rolling window size we use observations from July 1997 to December 1999 (30 observations) as the first rolling window to estimate possible unidirectional G-causality temporary cointegration for all the pairs (see appendix 1). To do that we do the following: We test cointegration by Johansen's method in that rolling window sample for example for USA as a leader (Y) and Finland as a follower (X) as one pair. If we find p-value for trace test $<$ 0.070, then the series are temporarily cointegrated and we can proceed to next stage which is the likelihood ratio tests for the long run weak exogeneity. If we do not find required p-value for cointegration or for weak exogeneity, we move to the next pair.) If we find that the boundaries are satisfied we do the maximum likelihood estimation of equations 3.8 & 3.8' for the current rolling window. We choose the smallest AIC to lag structure for 3.8 & 3.8 and we study whether p-value of α_x the ECM-variable satisfies the chosen criteria. In addition, we do diagnostic testing of the model. Then if all criteria are passed, we can make a forecast about the direction of change for January 2000 for the follower index and apply the trading rule on the basis of this forecast as presented earlier. Next we apply the same procedure for the rest of the pairs. After that we move to the next rolling window period (August 1997 – January 2000) (again 30 observations, because we have the rolling window) to forecast February 2000 and test all the pairs for possible forecast to February 2000, and so on until all 7047 possible forecastable positions have been tested.

3.7 IDENTIFYING INVESTOR'S SPECIFIC BIASED EXPECTATIONS

Shefrin (2000) notes that behavioral finance is everywhere that people make financial decisions. He continues that heuristic driven biases and frame dependence cause prices to stray from fundamental values, which leads to market inefficiency. The main biases are overconfidence, excessive optimism, representativeness, conservatism, availability bias, recency bias, illusion of validity, ambiguity aversion and gambler's fallacy bias. The main phenomena due to frame dependence are loss aversion, mental accounting, hedonic editing, regret aversion and myopic loss aversion. The two building blocks of behavioral finance are cognitive psychology (how people think) and the limits of arbitrage (when markets will be inefficient) (Ritter 2003).

Shefrin (2000) notes that market inefficiency does not necessarily mean economic profits opportunities. Shleifer & Vishny (1997) studied the limits of arbitrage and the impact of noise traders. They warn that noise traders can push prices away from the true value longer than it is tolerable for the arbitrageurs.

People herd when they do what others are doing rather than using their information (Banerjee 1992).There can be at least four kinds of herding in the financial market: simultaneous rational herding, simultaneous irrational herding, rational herding with a lead-lag and irrational herding with a lead-lag.

We propose that irrational herding with a lead-lag phenomenon in the asset prices may be caused by heuristic driven biases like illusion of validity (Kahneman & Tversky 1974) and recency bias (Murdock 1962). and overconfidence (Oskamp 1965),

If investors are overconfident about their ability or knowledge in difficult problems, they suffer from overconfidence bias (Oskamp 1965). The illusion of validity means that people make predictions by selecting a particular outcome that is the most representative of the input (Kahneman & Tversky 1974). Recency bias (Murdock 1962) implies that people tend to attach too much importance to recent experience.

Cipriani & Guarino (2008) show that in a two-asset economy herd behavior can generate long-lasting misalignments between prices and fundamentals. We demonstrate that in the two-asset economy similar to Cipriani & Guarino (with some further assumptions), where the investors have standard rational preferences but they may have biased beliefs, specific behavioral biases are sufficient conditions for existence of irrational herding with a lead-lag.

Daniel, Hirshleifer & Subrahmanyam (1998) argue that behavioral biases are increased when there is more uncertainty. They continue that return predictability should be stronger in firms with greater uncertainty because investors tend to be more overconfident when firms' businesses are hard to value. Zhang (2006) finds that returns are more predictable when investors face growing information uncertainty. Hirshleifer & Teoh (2003) argue that the tendency to herd might be greater when the private information that individuals receive is hard to process. They continue that if there are multiple dimensions of uncertainty, then something akin to a cascade may occur. Hirshleifer & Teoh (2003) define cascades as Banerjee defines herding. Shiller (2000) argues that even rational investors may participate to herding behavior when they take into account the judgements of others even when they know that everyone else is behaving in a herdlike manner. He continues that herding behavior produces group behavior, which is irrational.

We assume that investors have standard rational preferences but they may have biased beliefs. We argue that when investors who invest to the minor investing countries receive ambiguous information concerning their assets, they may become confused in the growing uncertainty. This may lead to a phenomenon where some investors take role models from large financial centers. In other words, there is herding behavior with a lead-lag effect. This phenomenon can be characterized as reputation based-herding in the spirit of Scharfstein and Stein (1990). We argue that herding may start because investors become more risk averters, which may act as a trigger for this particular herding behavior. If this herding behavior continues long and strong enough, the phenomenon may lead to temporary relative dependency with a lead-lag between particular assets. This may happen because investors suffer from an illusion of validity. There is a possibility that they are overconfident about the success of their current investing behavior as proposed by Daniel, Hirshleifer &

Subrahmanyam (1998). Therefore, their actions may affect a recency bias. We argue that these biases may lead to a temporary lead-lag effect between particular relative asset prices.

The efficient market hypothesis predicts that comovements in prices reflect comovements in fundamental values. Furthermore, it predicts that the comovements (that are due to global shocks) should happen simultaneously. However, the fundamentals may create herding behavior with a lead-lag, which leads to temporary cointegration with a lead-lag effect among assets, but this phenomenon should be identifiable with factor models (which take account of those fundamentals). That is, for some reason, economic fundamentals have a leader and a follower factor. These phenomena can be defined as rational herding. Shleifer (2000) and Barberis, Shleifer & Wurgler (2005) argue that comovements of fundamentally unrelated asset prices can be taken as evidence of the influence of investors' sentiment on asset prices. This phenomenon can be defined as irrational herding. In other words, there can be at least four types herding in international asset markets:

1. Rational simultaneous comovements: These are caused by the comovements of the fundamentals (global or country specific). We may identify this with bidirectional temporary cointegration and with the factor models. This can be characterized as rational herding.

2. Irrational simultaneous comovements: These are due to the comovements of fundamentally unrelated asset prices, which move simultaneously. We may identify this with bidirectional temporary cointegration. This is irrational herding.

3. Rational comovements with a lead-lag effect: These are caused by the comovements of the fundamentals with a lead-lag effect. Thus, the fundamentals have a leader and a follower. We may identify this with the unidirectional temporary cointegration model and with the factor models. This is rational herding with a lead-lag.

4. Irrational comovements with a lead-lag effect: These are the comovements of fundamentally unrelated asset prices where is a leader asset and a follower asset.

We may identify this with the unidirectional temporary cointegration model. This is irrational herding with a lead-lag.

Summary 2. Unidirectional G-causality temporary cointegration is a consistent with herding with a lead-lag effect, (rational and irrational).

The conclusion for the temporary cointegration with unidirectional Granger causality in asset prices, we propose the following.

Proposition 1. Asset prices may start to deviate from their fundamental values or from their different kinds of sentiment pricing in this particular case because of the possibility of investors' specific behavioral biases in times of growing uncertainty. The risk aversion may act as a trigger. With recency bias, the illusion of validity and overconfidence this may lead to irrational unidirectional temporary cointegration between particular assets. Thus, in that period specific sentiment may weigh more than the fundamentals in investors' decisions. This situation lasts until debiasing may start to take effect and the fundamentals or a different kind of sentiment may start again to weigh more than this specific sentiment.

3.7.1 DEMONSTRATION OF PROPOSITION 1

Dasgupta & Prat (2008) show that in a one-asset economy, when there are career-concerned (one example of reputation-based herding) institutional traders as market makers, prices never converge to true value even after infinite sequence of trades. Cipriani & Guarino (2008) show that in the two-asset (correlated or independent) economy specific sequences of trades generate informational cascades and long-lasting misalignments between prices and fundamentals. They assume that investors have heterogeneous private true values of assets. Cipriani & Guarino argue that in their economy informational cascades can push the price of the other asset away from its fundamental value even in the long run, because investors expect to have gains from trade. We demonstrate that in the two-asset economy close to Cipriani & Guarino, where the investors have rational preferences but they may have biased beliefs, specific behavioral biases are sufficient conditions for existence of irrational herding with a lead-lag.

3.7.1.1 MODEL

In the model we follow the assumptions of Cipriani & Guarino (2008). **The following assumptions are the same:** Let us assume that investors have standard rational preferences. Let us assume a two-asset economy (stock X and stock Y), infinite time horizon and for simplicity assume that risk free rate is zero. There are two kinds of investors; informed and uninformed traders. The investors can buy, sell or decide not to trade. In this economy, investors' expected true value of asset (stock X and stock Y) is the sum of expected future cash flows from the asset discounted with subjective rate r. Thus, investors have heterogeneous private expected true values of the assets. Let us assume that private expected true value of the asset is greater than 0. Let us assume that the traders of asset X (Y) at the time t observe only the history of prices Y and X. Let us assume that investors are pricing their assets at all times. Let us assume that expected true values of the asset (X and Y) follows martingale process in time. Thus, they are independent sequentially.

The differences: Cipriani & Guarino assume that investors act in an exogenously determined sequential order and they interact with a market maker. At each time t an investor can exchange the asset with the market maker. In their model the market maker sets the prices at which a trader can buy or sell the asset. In our model there is no market maker but we define that the trading is done with open auction and every investor can participate in it, but only one auction at same time. Thus, common priors (Aumann, 1987 & 1998) determine equilibrium (closing price) at the time t. Common priors means that investors with the same beliefs and discount factors have the same true value of the asset at time t. In addition, we define in the spirit of Cipriani & Guarino that investor who does not participate in the auction, recognizes only closing price of the assets. That is even if investor A holds both asset X and Y, she can only participate trading auction one asset at the time, therefore she participates in to the auction of asset X, she recognize only closed price of Y (past prices of Y). Thus, because of the auction style we assume that closing prices are always equilibrium and thus investors themselves construct the market maker.

Our further assumptions are: Cipriani & Guarino assume that investors' expected true value of the assets X and Y at the time t can take only two values, lower than market maker's price or higher than market maker's price. We assume further (because of the auction, and every investor can participate to it) that investors' expected true values of the assets at the time t construct continuous distributions (for example $\alpha P \leq Y_1 \leq \beta P$, where P is price of a share Y at the time 1, $0<\alpha<1$ and $\beta>1$).

In addition to Cipriani & Guarino model we define that there are two groups of investors. The group FIN are investors who are the holders of stock X at the time t. The group USA are investors who are the holders of stock Y at the time t. Let also assume that the group USA is relatively larger than the group FIN and the group USA has **reputation** to be more informed than the group FIN. Both groups are pricing their stock under subjective belief about future cash flows which are discounted with some subjectively believed correct discount rate. Furthermore, to Cipriani & Guarino model we define the following: Short selling is allowed. We assume that investors have standard rational preferences but they may have biased beliefs and investors are risk averters.

TIME = 0

Let us assume that at the time 0 common priors produce prices of stocks Y and X to

$Y_0 = P$
$X_0 = P/\varphi$.

TIME = 1

Let us assume that at the time 1 growing multidimensional uncertainty occurs in the markets. The dimensions are: the effect of a shock to the asset value, the existence of a shock and the quality of traders' information (Avery & Zemsky 1998). In other words, there is a growing information uncertainty and less informed investors may become more confused in the growing uncertainty. This may lead to uncertainty of how ongoing shocks affect the true values of stocks X and Y. That is, expected subjective true value of Y for the group USA at the time 1 is

$\alpha P \leq Y_1 \leq \beta P$,

where $0<\alpha<1$ and $\beta>1$ and expected subjective true value of X at the time 1 for the FIN is

$\gamma(P/\varphi) \leq X_1 \leq \kappa(P/\varphi)$,

where $\gamma<\alpha$ and $\kappa>\beta$. Thus, the FIN is more confused in the growing multidimensional uncertainty than the USA. Let us assume that closing prices (common priors produce) for Y and X at the time 1 are

$Y_1 = aP$

and

$X_1 = bP/\varphi$,

where $\alpha<a<\beta$ and $\gamma<b<\kappa$.

TIME=2

Let us assume that at the period 2 multidimensional uncertainty continues and increases in the markets. That is, expected subjective true value of Y for the group USA at the time 2 is

$(\alpha-\lambda)P \leq Y_2 \leq (\beta+\lambda)P$

and expected subjective true value of X for the group FIN is

$(\gamma -\lambda) (P/\varphi) \leq X_2 \leq (\kappa +\lambda) (P/\varphi)$,

where $0<\lambda<1$. There is a possibility that because of the risk aversion less informed investors may take temporary hint for their investing decisions from the group that has a reputation to be more informed. That is the majority of the FIN may start reputation-based herding from the USA. Because the FIN recognizes only past prices of Y they may decide to make offers until the price for $X_2=a(P/\varphi)$ is reached. This may happen because the majority of the FIN has a belief that current shocks are global and true value of X is the same as it was relatively to Y at the time 0. Thus the majority of the FIN believes that true value of X= $a(P/\varphi)$. Let us assume that closing price for Y (common priors produce) at the time 2 is

$$Y_2 = \eta aP.$$

where $0<\alpha<\eta<1$. Because of the reputation-based herding and because of the fact that the investors who participate to the auction of X recognize only past prices of Y, the closing price for X (common priors produce) at the time 2 is

$$X_2 = a(P/\varphi).$$

TIME=3

Let us assume that multidimensional uncertainty continues at the time-step 3. The majority of FIN may believe that reputation-based herding works well because the closing price of X at the time 2 was $a(P/\varphi)$. That is just what they have anticipated. That is why they may continue to herd. Thus, they may suffer from **illusion of validity** and **recency bias**. The majority of the FIN may have subjective belief that true value of X is now $\eta a(P/\varphi)$. Even if they may have private information about different kind of true value, they choose to ignore it. They are maybe overconfident about their current (herding) pricing method. Thus, they suffer from **overconfidence** bias as proposed by Daniel, Hirshleifer & Subrahmanyam (1998). Thus, they may learn to herd because of behavioral biases (illusion of validity and recency bias).

Let the expected subjective true value of Y for the USA at the time 3 be

$$(\alpha-\lambda)P \leq Y_3 \leq (\beta+\lambda)P$$

and because of possibility that behavioral biases affects for expected subjective true value of X for the FIN, it is then

$(\alpha-\lambda)P/\varphi \leq X_3 \leq (\beta+\lambda)P/\varphi.$

Let assume that closing prices for Y (common priors produce) at the time 3 is

$Y_3 = \upsilon\eta aP$

where $0<\alpha<\upsilon<1$ and maybe because of the reputation-based herding and because of the investors who participate in the auction of X recognize only past price of Y, the closing price for X (common priors produce) at the time 3 is

$X_3 = \eta\alpha\ (P/\varphi).$

Thus, X_t and Y_t have developed the same stochastic trend. They have unidirectional temporary cointegration where Y_t is a leader asset and X_t is a follower. The generalization for this phenomenon is

$X_t = 1/\varphi Y_{t-1},$

Uninformed investors may learn to herd, because of behavioral biases. If informed investors have private information about the true value of X and the true value is different from the herding true value, and they choose to ignore it, behavioral biases are sufficient condition for herding with a lead-lag. Thus, herding is irrational, because we assume standard rational preferences and that investors are pricing their assets at all times. Learning from the herd is irrational, because rational investor cannot have a rational belief that irrational herding continues. This can be seen from the fact that, in this model, if informed investor A considers to follow the herd until time t+h, because she expects to have gains from it (that is, if the herding behavior continues until t+h), she has to take into account that informed investor B can decide to herd until time t+h-1. Then, investor B considers to follow the herd until t+h-1, but she has to take account that informed investor C can decide to herd until time t+h-2 and so on. Therefore, in the optimum (because every investor can participate in the auction) informed investor uses her private information at the time t. **Thus, rational preferences with unbiased beliefs imply that irrational herding will end immediately and price of the asset will go back to its expected fundamental value immediately. This happens, because of the previous reasons and the fact that unbiased informed investors do not know the quantity of uninformed investors at the time t.** Previous reasons imply that if we assume that multidimensional uncertainty continues forever and if informed investors who recognize behavioral biases will not take advantage from their private information because of the limits of arbitrage, prices never go back to fundamental value. Thus, under unidirectional G-causality temporary cointegration, $X=Y/\varphi$ is long run equilibrium and the price of the asset X can be predictable. Unidirectional G-causality temporary cointegration may last until multidimensional uncertainty ends in the market and the debiasing may start to take effect or the miss pricing is large enough or a different kind of sentiment may start to weigh more than this specific sentiment.

3.7.2 CASE STUDY: UK AS A LEADER AND DENMARK AS A FOLLOWER

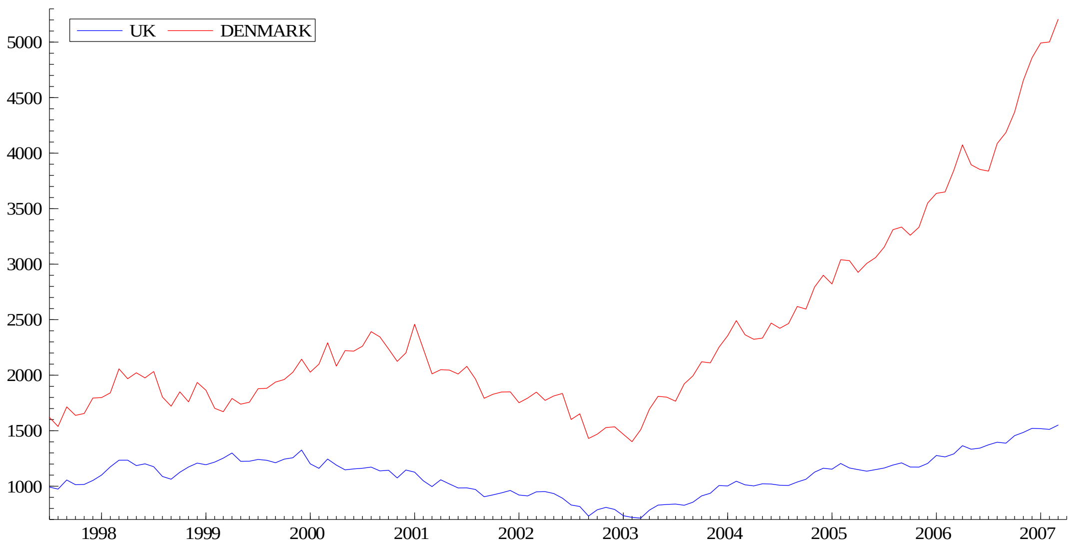

Figure 1. Price series in UK and Denmark from July 1997 to March 2007.

Figure (1) presents price series of UK and Denmark MSCI indices from July 1997 to March 2007. According to our tests and testing criteria there was a temporary relative lead-lag effect between these assets from May 2000 to March 2003, which we can identify with our technique in October 2002 under the real time principle. We argue that when the technology assets bubble began to burst at the beginning of 2000 people who invested in the Danish stocks may have received ambiguous information about their assets in Denmark, because of that incident. They might have become worried about how long and how severe the decline would be. Figure (2) shows closer look before May 2000.

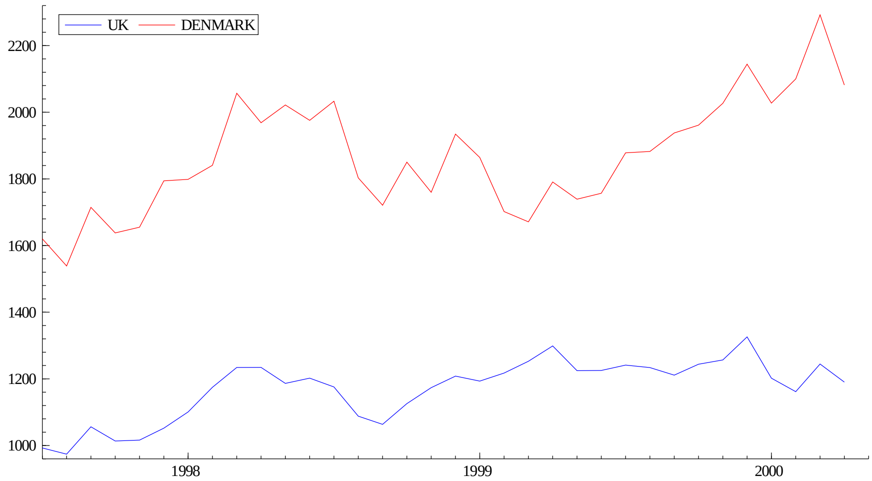

Figure 2. Price series in UK and Denmark from July 1997 to April 2000.

According to our tests, we cannot identify temporary cointegration between (Figure 2) these price series before January 2000. However, information uncertainty may have been growing among those holding Danish stocks during the bursting of the technology bubble at the beginning of 2000.

We argue that because of information ambiguity and risk aversion among holders of Danish stocks, they may have begun to take temporary hints for their investing decisions from those investors who invested in UK stocks. Therefore, reputation based-herding may have developed such that investors in Danish stocks herd from UK investors for directions about relative future returns starting at the beginning of 2000. Because of short-term data showing that herding behavior apparently worked well among investors in Danish stocks, they may have became overconfident about the validity of herding. In other words, at this point, investors may have suffered from an illusion of validity and overconfident biases. They may have put too much weight on recent experience. This stems from recency bias. Figure (3) shows the time line between UK and Denmark from May 2000 to March 2003, which is the temporary cointegration with the unidirectional Granger causality period between UK and Denmark stock indices in this sample.

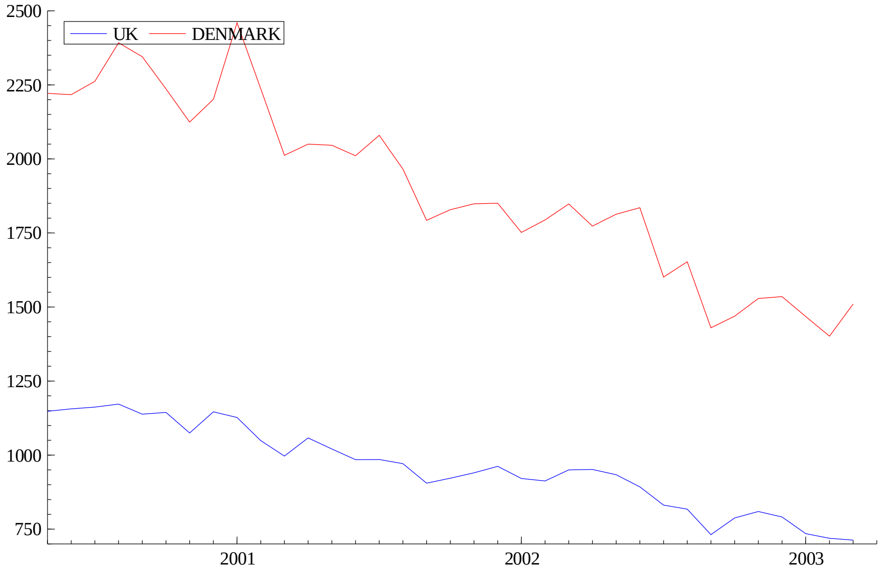

Figure 3. Price series of UK and Denmark during the temporary cointegration from May 2000 to March 2003.

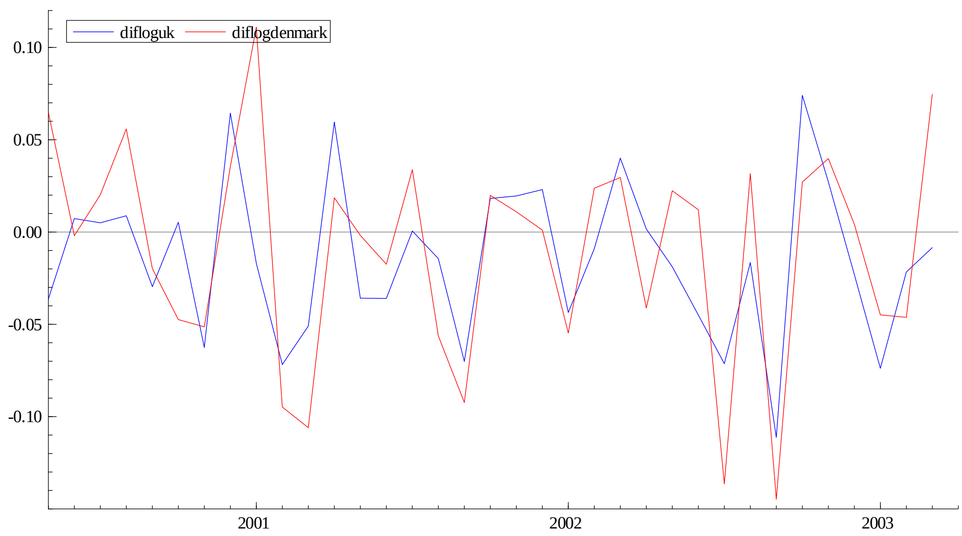

Figure 4. Changes in logarithmic prices series of UK and Denmark during the temporary cointegration from May 2000 to March 2003.

Figure 4 shows the changes in the logarithmic prices series of UK and Denmark during the temporary cointegration from May 2000 to March 2003. Changes in logDenmark (blue line) series seem slightly more volatile than changes in logUK (red line) series, but it appears that the changes in logUK and logDenmark go hand in hand. According to our Granger causality test in this period lead-lag effect goes from UK to Denmark.

After March 2003 something has happened, because (according to our tests) temporary cointegration between UK and Denmark vanished. Figure (5) shows the price series for UK and Denmark indices from April 2003 to March 2007. We argue that after March 2003 the illusion of validity and recency bias may have had a diminishing effect on their investment decisions among holders of Danish stocks in this particular case. From this we can conclude that the sentiment which may have created temporary cointegration with unidirectional Granger causality between UK and Denmark assets has vanished and a different kind of sentiment or the fundamentals again weigh among holders of Danish stocks.

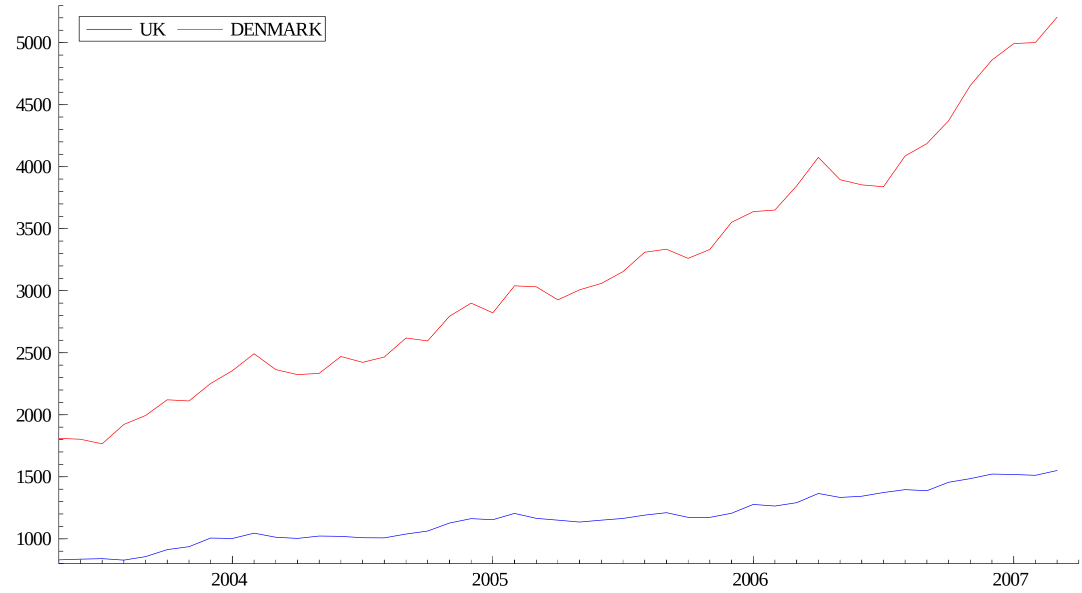

Figure 5. Price series for UK and Denmark from April 2003 to March 2007.

4. TESTING RETURN PREDICTABILITY WITH STOCK INDICES

In this section, we present details for the trading rule that we use to test return predictability with unidirectional G-causality temporary cointegration technique.

DATA

Timmermann and Granger (2004) argue that the latest forecasting techniques may have a "honeymoon" period before their use becomes more widespread and then economic profits will rapidly disappear through these investors' transactions. To deal with this phenomenon, it is easy to conclude that when return predictability is tested, it is important to use the very latest data available. In this paper, we therefore use the testing period data from January 2000 – March 2007.

Furthermore, Pesaran & Timmermann (1995) warn: "In particular, it is important that, as far as possible, rules for prediction of stock returns are formulated and estimated without the benefit of hindsight." For this reason all the trading rule tests have been made in "real time". This means, for example, that when the predictability of data until 31. December 1999 is tested, we use only information from the data up to that particular date. When we identify temporary cointegration with unidirectional G-causality, **we open our position at the next trading day**. It means, for example, if the technique suggests that January 2000 is forecastable in Hong Kong index, we use only data up to December 1999 to make that forecast and we open the position in January 2, 2000.

Data is MSCI (Morgan Stanley Capital International) country indices as monthly data. These indices describe the developments of individual country stock prices where stock weights have been defined by Morgan Stanley. We chose these indices as the object of the tests, because we assumed that they would include less noise than individual stocks. Monthly data was developed such that every month is represented by the last official quote. We chose a month as a time step because we assume that in this way it is possible to detect same trend of variables more easily

In this study MSCI-data includes all country indices which currently belong to MSCI WORLD index. They are UK, USA, Japan, Germany, Holland, Belgium, Italia, Hong Kong, Spain, Finland, Norway, Sweden, Denmark, Austria, Greece, Ireland, Singapore, New Zealand, Australia, Canada, Portugal, France and Switzerland. The testing period is 87 months from the January 2000 to March 2007.

4.2 TRADING RULE

To test a trading rule, we have to develop a trading rule simple enough to be coded in a computer that can repeat it automatically in real time. Subrahmanyam (2008) argues that ex-ante tests take care of the data mining problem.

According, for example, to Fama (1991) and Timmermann & Granger (2004) the best choice for the normal return is observed return during the testing period. In this study the investment space is MSCI indices belonging to the MSCI WORLD index. Therefore as normal returns the MSCI WORLD index was selected which takes account of dividends under the testing period.

Shleifer & Vishny (1997) studied the limits of arbitrage and the impact of noise traders. They warned that the noise traders can push prices away from the true value longer than it is tolerable for the arbitrageurs. We therefore restrict our open positions to one month only.

In order to find predictive patterns from the series tested, we have to make possible lead-lag pairs from the country indices. We assume that USA, UK and Switzerland are leader countries in the finance world and are possibly followed by other countries. The first lead-lag pairs are

Leader	Follower
USA UK Switzerland	Finland Norway Sweden Denmark Austria Belgium Italy Spain Greece Ireland Portugal France Holland Singapore Australia Canada Hong Kong New Zealand

For these pairs the months under testing during time period amount to 3 * 18 * 87 = 4698. Additionally, a potential own leader country is chosen for three continents. Germany is chosen for Europe, Japan for Asia and Hong Kong for America. This way the second pair setting is:

Leader	Follower
Germany	Finland Norway Sweden Denmark Austria Holland Belgium Italy Spain Greece Ireland France Portugal
Japan	Singapore Hong Kong Australia New Zealand
Hong Kong	Canada

For these pairs the months under testing amount to 1 * 13 * 87 + 1 * 4 * 87 + 1 * 1* 87 = 1566. Furthermore, world index is tested with leader indices (where in these cases leader indices become potential followers for the world index) and USA is tested with leader indices:

Leader	Follower
World index	USA UK Switzerland Japan Germany

Leader	Follower
USA	UK Switzerland Japan Germany

For these pairs the months under testing amount to 1 * 5 * 87 + 1 * 4 * 87 = 783. Total months under testing amount to 7047. All tested lead-lag pairs (81) can be found from appendix (2).

It would appear characteristic of the method that it detects special cases from the investment space where statistical arbitrage opportunities might be found. Therefore in our method all assets are invested in acceptable (targets whose trading rule criteria have been accepted) indices such that assets are divided equally between the indices in question each month. If the method cannot find any acceptable index to invest in for some month, it can be interpreted that the method advises investing in to risk free assets. In this study risk free asset is US one-month treasure bill.

In order to device maximum benefit from the forecastability of the direction of the price changes, short selling is permitted in this study. Then the asset in question is borrowed from the financial market.

When the technique suggests that the index in question is going to rise in the next month, that index is bought in the **following day**. Then position is open. The position is closed when this forecastable month end for the closing price.

By contrast, when the trading rule suggests that there will be negative returns for the index in question, that index is sold in the **following day**. So then, the position is open. The position is closed when this particular month end for the closing price. Therefore, the maximum short selling period is one month.

In other words, the trading rule is the following.

1) In every time step (one month) the position-opening period is **one month minus one day** in the testing period.

2) If the technique forecasts (for some month) that there is not going to be one-step ahead temporary cointegration with unidirectional Granger causality between MSCI indices then all assets are invested in the risk free assets (US one-month treasure bill) in that period.

3) If the technique forecasts (for some month) that there are **some** MSCI pairs with temporary cointegration with unidirectional Granger causality then all assets are invested at the next trading day in the **identified follower** from the identified MSCI pairs by dividing all assets **equally** amount all identified follower-indices for that period.

4) If the technique forecasts (for some month) that there is **one** MSCI pair to invest then all assets will be invested at the next trading day in the **follower** MSCI index.

5) If the technique finds (for some month) at least two MSCI pairs to invest in, and the forecast for one index is that the next step is going to be a positive change and the other is negative change, then the technique suggests that an statistical arbitrage opportunity exists. As technique proposes then we buy the positive (forecasted to be positive change) index and sell (forecasted to be negative change) the negative index.

6) If the technique forecasts (for some month) that the next step is going to be a negative price change then we **sell** that asset at the next trading day.

7) If the technique forecasts (for some month) that the next step is going to be a positive price change then we **buy** that asset at the next trading day.

4.3 TESTING THE TRADING RULE

From the testing period January 2000 – March 2007, the technique finds 44 forecastable months. Total of **73 positions** are opened. That is 1.0 % from possible position openings (73 / 7047). For the remaining months (43) assets are invested in US one-month treasure bills. Table (1) shows the months, which the trading rule technique suggests to be forecastable and number of positions opened. In these forecasts, the hit ratio is 80.8 %. The hit happens when the sign of forecasted price change is the same as observed price change. The hit ratio data can be found from the appendix (3). We assume that the transaction costs are 0.5 % per transaction and the short selling borrowing rate is 12 % in annual terms. For example, as we sell short Hong Kong index in February 2007 we get the following return because our prediction is correct: +4.26 % − (1% + 0.5% + 0.5%) = +2.26%.

Table 1. Forecastable months and number of opened positions.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	total
2000			1	1				1	1				**4**
2001				1	1				2	1	1	1	**7**
2002						2	1	2	2	2	2		**11**
2003	2	2	3	1				1	1				**10**
2004				1	2	2	2	1	2	1	4	3	**18**
2005	2	1		3		1	5					1	**13**
2006	1		1		1					1			**4**
2007	2	1	3										**6**
total	**7**	**4**	**8**	**7**	**4**	**5**	**8**	**5**	**8**	**5**	**7**	**5**	**73**

Figure (9) shows the trading rule returns with costs and world index returns series (benchmark). As we can see, the difference between these two series is remarkable. In particular, the trading rule series includes significantly less months of negative returns.

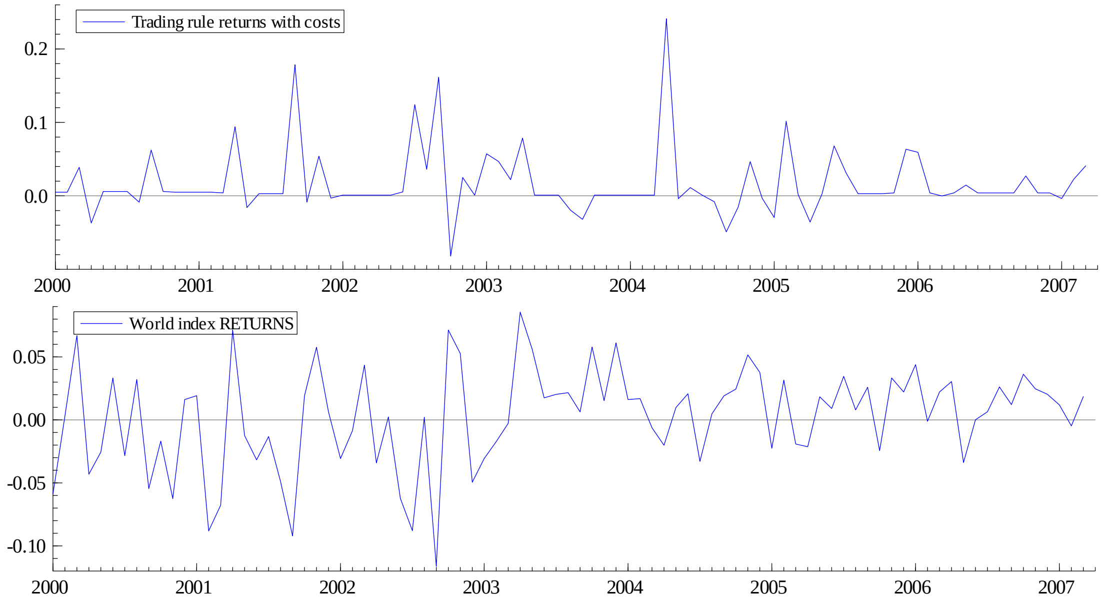

Figure 9. Trading rule returns and benchmark returns during out-of-sample testing period January 2000 – March 2007.

Table (2) shows the descriptive statistics of these series. We can see that the distribution of the trading rule returns is skewed to the right (+2.35). On the other hand, the benchmark returns are skewed to the left (-0.57). Excess kurtosis (7.67) and skewness imply that the trading rule returns are concentrated on near zero and to large positive returns. For the benchmark, kurtosis and skewness imply that returns are close to normal distribution except larger distribution to negative side. However, according to a small sample adjusted Jarque & Bera test, the benchmark returns follow normal distribution with p-value (0.08). Figure (10) reveals the same thing graphically.

Table 2. Descriptive statistics of the trading rule returns with costs and benchmark.

	Trading rule returns	World index returns
Observations	87	87
Mean	0.017	0.002
Standard deviation	0.046	0.040
Skewness	2.346	-0.570
Excess Kurtosis	7.670	0.281
Minimum	-0.082	-0.116
Maximum	0.241	0.086
Number of negative returns	17	34
Average negative return	-0.021	-0.037
Average positive return	0.027	0.028

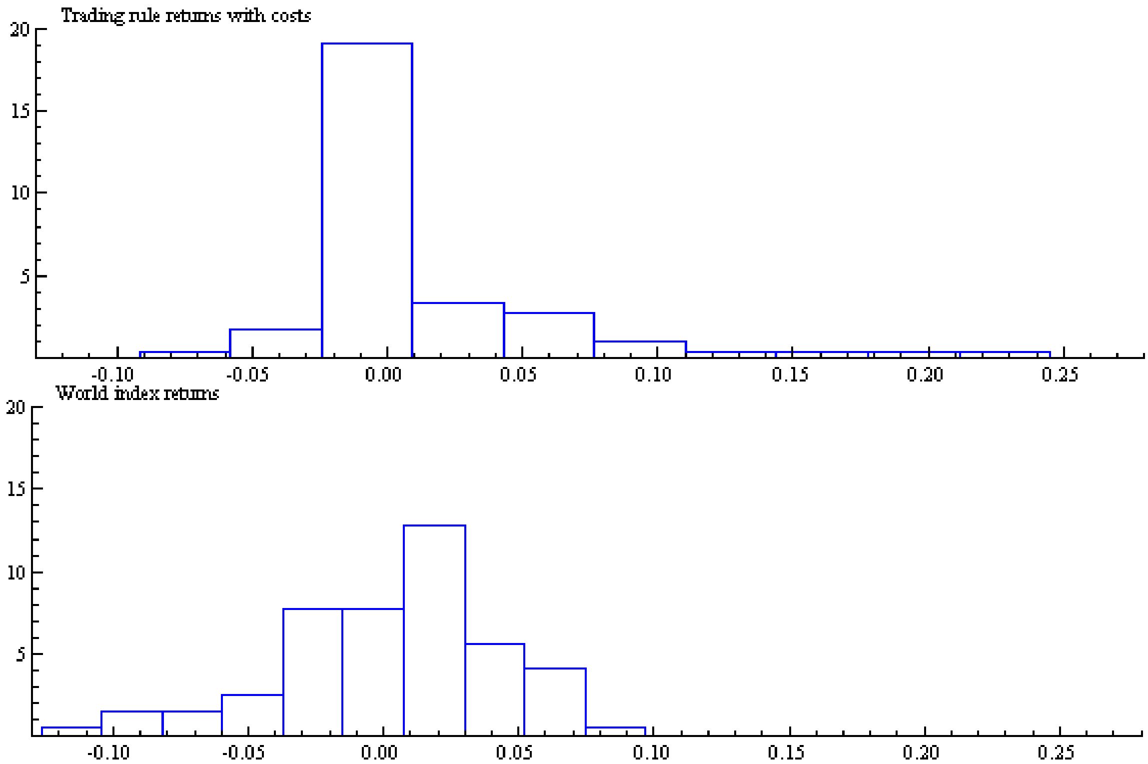

Figure 10. Distributions of the trading rule returns and the benchmark.

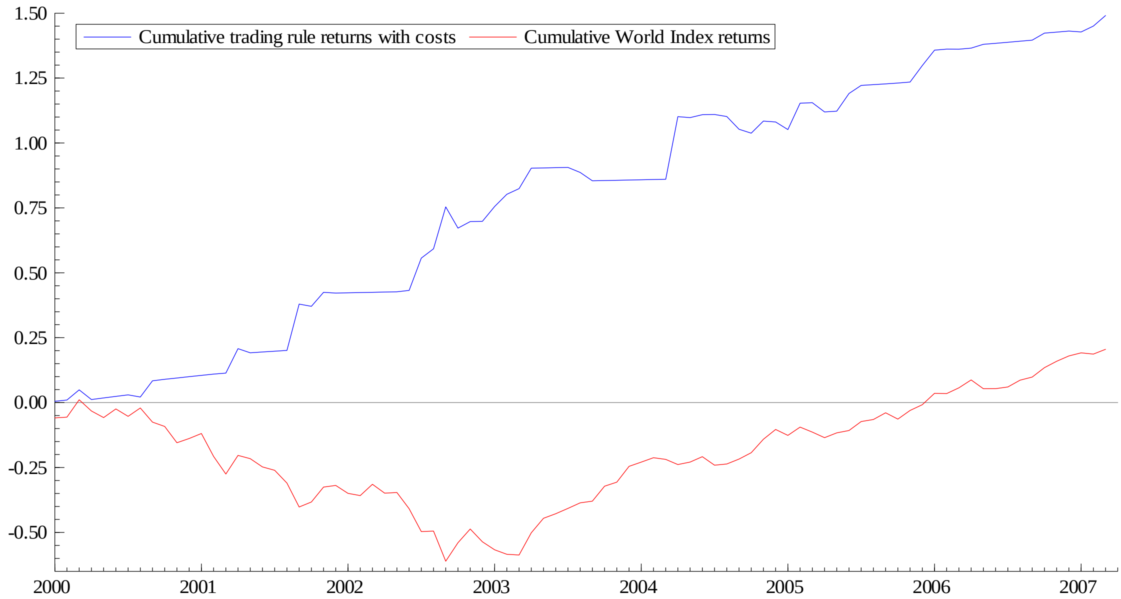

Figure 11. Cumulative returns of the trading rule and the benchmark.

Figure (11) shows cumulative returns. It looks like major part of the difference between the trading rule returns and world index returns comes from the bear market conditions during 2000 – 2003. In addition, Figure (11) may suggest that trading rule produces consistent profits regardless of market conditions. Thus, the trading rule strategy can be considered as market-neutral hedge fund strategy. Cumulative returns over the whole testing period for the trading rule are +149.1 % and for the benchmark +20.5 %.

4.3.1 STATISTICAL AND ECONOMIC SIGNIFICANCE OF THE TRADING RULE RETURNS

In this section, we present four tests where we examine statistical and economic significance of the trading rule returns. According to EMH the hit ratio of the trading rule forecasts should be 50 %. Thus, if 50 % were exceeded with statistical significance it would be an evidence against EMH. The performance of the trading rule returns is examined with Jensen's α. Riskiness of the trading rule returns is studied with the Value at Risk analysis. As we note in the previous chapter, the trading rule strategy is a market-neutral hedge fund strategy. According to Getmansky, Lo and Makarov (2004) the market-neutral-strategy hedge funds may suffer significant serial correlation in the returns because of the liquidity exposure.

Thus we examine serial correlation of the trading rule returns to rule out strategy's liquidity exposure as proposed by Lo (2008).

We take 73 bets (Table 1) during the trading rule test. The hit ratio for the forecasts (positions) is 80.8 %. Pesaran & Timmermann (1992) propose a market-timing (PT) test. Let P be the hit ratio, i.e. the proportion of cases where binary variables Y_t (observed outcomes, 1 or 0) and X_t (forecasted outcomes, 1 or 0) fall in the same category (have the same sign), while P* is the hit ratio expected under the null of independence between Y_t and X_t (EMH forecasts that P should be 50 %). Pesaran & Timmermann continue that the PT test statistic is

$$\mathrm{PT} = \frac{P - P^*}{[V(P) - V(P^*)]^{1/2}} \qquad \sim \mathrm{N}(0,1)$$

where

$$\mathrm{V(P)} = \mathrm{T}^{-1}\mathrm{P}^*(1 - \mathrm{P}^*),$$

$$\mathrm{V(P^*)} = \mathrm{T}^{-1}(2(\underline{Y} - 1)^2\, \underline{X}(1 - \underline{X}) + \mathrm{T}^{-1}(2\underline{X} - 1)^2\, \underline{Y}(1 - \underline{Y}) + 4\mathrm{T}^{-2}\, \underline{YX}(1 - \underline{Y})(1 - \underline{X}),$$

T is observations, $\underline{X}$ is an average of binary variable X, and $\underline{Y}$ is an average of binary variable Y. The PT-test confirms that the percentage of correct signs is statistically significant. P-value is < 0.0001. However, this test assumes serially independent outcomes. We have overlapping data. In such cases, it may become important to allow serial correlation among discrete variables. Pesaran & Timmermann (2006) propose a regression approach test, which allows serial dependence. They argue that PT statistics can be approximated by the t-ratio of the coefficient β in the ordinary least square regression of Y_t (observed outcomes, 1 or 0) on X_t (forecasted, 1 or 0) and an intercept α. The equation becomes

$$Y_t = \alpha + \beta X_t + u_t,$$

where u_t could be serially correlated and heteroscedastic. The t-ratio of β is given by

$$t = \frac{r\sqrt{T-2}}{\sqrt{1-r^2}}$$

where r is correlation coefficient between binary variables Y_t and X_t. Pesaran & Timmermann argue that

$$PT \approx \frac{\sqrt{T}}{\sqrt{S_{yy}S_{xx}}}\frac{S_{yx}}{} = \sqrt{T}\, r,$$

where S_{xx} is the variance of X, S_{yy} is the variance of Y and S_{yx} is the covariance between Y and X. For more information about the methodology for this test, see Pesaran & Timmermann (2006). P-value for the hit ratio from this test is < 0.0001 in our trading rule outcomes.

Jensen (1969) proposes a risk adjusted performance measure (Jensen's α), which can be calculated by CAPM:

$$[R_{it} - R_{ft}] = \alpha_i + \beta[R_{mt} - R_{ft}], \tag{4.1}$$

where R_{it} is the trading rule return, R_{ft} is the risk free return (US one-month treasure bill) and R_{mt} (World index return) is the market return. According to CAPM the return for the trading rules should be $\beta[R_{mt} - R_{ft}]$. Therefore, if we have a positive and statistically significant α_i it suggests that an economic profit is gained. Table (3) shows the statistical results where the trading rule returns (87 observations) are modeled by CAPM (equation 4.1).

Table 3. Statistical results for modeling the trading rule returns with costs minus risk free returns by OLS.

	Coefficient	HSCE	t-value	t-prob
α_i	0.0143	0.0048	2.96	0.004
β_i	-0.236	0.1879	-1.26	0.210

.

The trading rule produces statistically and economically significant profits (α_i). We use heteroscedasticity consistent standard errors, because we observed heteroscedasticity in the residuals. Additionally, there is no autocorrelation in error term according to LM-tests.

The Value at Risk (VaR) provides an answer to the following question: with a given probability α (we choose $\alpha = 0.05$), what is my forecasted financial loss over a given time horizon? J.P. Morgan (1996) introduced the Riskmetrics method where the VaR is calculated as follows:

$$\sigma^2_t = \omega + (1 - \lambda)\varepsilon^2_{t-1} + \lambda \sigma^2_{t-1} \tag{4.2}$$

where σ^2 is conditional variance of the return, ω is the constant, ε^2 is the square of the return and λ is fixed in the basic form. The Riskmetrics model is IGARCH (Nelson 1990) model with fixed λ. Because we have the monthly data, we fixed λ at 0.97 (as it is a common practice in a monthly data). We use AIC criteria to choose the best model from the GARCH-family. For the trading rule returns the best model is GARCH (1,1) with asymmetric student distribution. Bollerslev (1986) introduced a GARCH model. For the world index returns the basic model (4.2) is the best according to AIC. Ex ante conditional VaR value for the trading rule one-step ahead is -0.005 and for the world index returns -0.056. It means that in the next period, there is a 95 % probability that losses will stay below 0.5 % (5.6 %) and a 5 % chance they break through it for the trading rule (for the benchmark) according to conditional VaR.

Figure (12) shows the Value at Risk graphic analysis for the trading rule returns and World index returns. The red line shows a level where no more than 5 % of the negative returns is exceeded and the dotted blue line shows a level where no more than 5 % of the positive returns is exceeded during the testing period (January 2000 – March 2007).

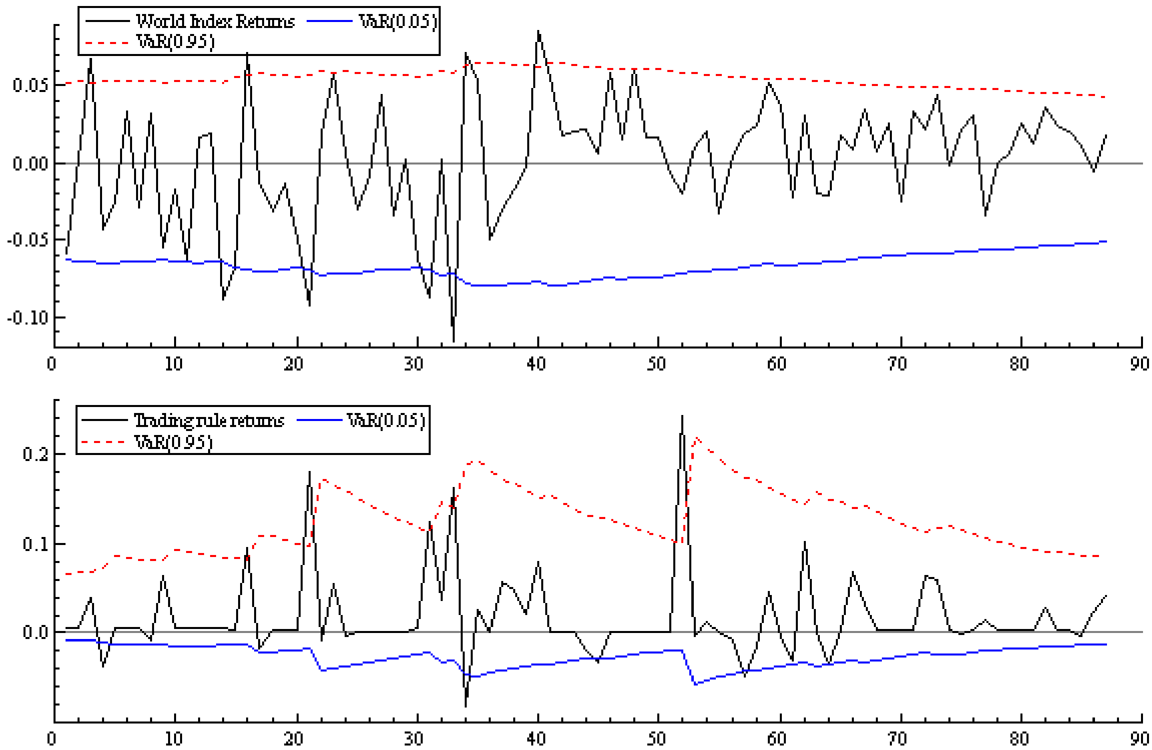

Figure 12. Value at Risk graphic analysis during the testing period.

Getmansky, Lo and Makarov (2004) and Lo (2008) note that especially the market-neutral-strategy hedge funds may have significant serial correlation in the returns. They argue that such correlation can yield substantial biases in the variances, betas and other performance statistics. Lo tests returns' serial correlation up to six lags. We test possible autocorrelation of returns with Ljung-Box test (1978) up to six lags. The p-value is 0.59 where the null hypothesis is that there is no serial correlation. Thus, we conclude that there is no autocorrelation in the trading rule returns.

5. DISCUSSION ON POSSIBLE REASONS FOR ECONOMIC PROFITS

5.1 RISK FACTORS MODEL TEST

There are at least two techniques to detect possible causes for the economic profits. One is to forecast (ex-ante) with another procedure (which reflects possible reason for the profits) the same observations as in the trading rule and compare the results. Another one is to determine (ex-post) the explanatory power of common risk factors that may produce the trading rule returns by multi risk factor model as originally proposed by Sharpe (1992). Further developments of this test to hedge funds analysis have been provided by Fung & Hsieh (1997), Ennis & Sebastian (2003) and Lo (2008) among others. We use the method and the risk factors as suggested by Lo. We construct OLS regression of the trading rule returns with costs (87 observations, January 2000 – March 2007) on the following six risk factors:

1) USD: the U.S. Dollar return;
2) BOND; return on the Vanguard Long-Term Bond Index
3) CREDIT; the spread between US long-term bond yield and the US short-term interest rate (Source OECD)
4) WORLD INDEX; the MSCI World Index return
5) CMDTY; return on the CBOE Gold Index
6) DVIX; the first difference of the end-of- month value of the CBOE Volatility Index.

Lo (2008) argues that these six risk factors correspond to basic sources of risk and of expected returns. USD reflects risk of currencies. BOND stems for the bond market risk. CREDIT reveals changes that are due to credit markets. We use the US credit market as a proxy for the credit market of the world. WORLD INDEX shows risk premium of the stock market. CMDTY reflects returns that are due to the developments of commodities prices. We use gold as a proxy for commodities. Finally, DVIX reflects returns that are due to changes of volatility.

According to Lo, ordinary least square estimation method provides a simple but useful decomposition of a return R_t into these components. Linear regression model becomes

Trading rule returns with $\text{costs}_t = \alpha + \beta_1 USD_t + \beta_2 BOND_t + \beta_3 CREDITt + \beta_4 WORLD\ INDEX_t + \beta_5 CMDTY_t + \beta_6 DVIX_t + \varepsilon_t$. (5.1)

Lo argues that the magnitude of trading rule alpha depends on how much of a trading rule's expected return is driven by common risk factors versus trading rule-specific alpha. He continues that this can be measured empirically. In other words, Lo argues that this characterization implies that there are two distinct sources of a trading rule's expected returns: beta exposures β_k multiplied by the risk premia associated with those exposures E[Risk Factor_{kt}] and trading rule-specific alpha α.

To reveal possible risk factors reasons for economic profits we estimate model (5.1) by OLS as suggested by Lo. Table (5) shows the statistical results where the trading rule returns are modeled by the risk factors model. We use heteroscedasticity consistent standard errors, because we find heteroscedasticy in the diagnostic testing. Additionally, there is no autocorrelation in the error term according to LM-test or ARCH-effect according to LM-test using squares. **According to statistical results (table 5), the economic profits are not due to risk factors that are included in the model** (5.1). We do not to imply that the trading rule returns' unique source of alpha is without risk. We argue that according to the estimation results (table 5) these risk factors (USD, BOND, CREDIT, WORLD INDEX, CMDTY & DVIX) are not source of economic profits.

TABLE 5. Statistical results for modeling the trading rule returns with costs (87 observations, January 2000 – March 2007) on six common risk factors by OLS.

	Coefficients	HCSE	t-value	t-prob
α	0.018	0.005	3.52	0.0007
USD	-0.196	0.204	-0.96	0.34
BOND	-0.457	0.284	-1.61	0.11
CREDIT	1.619	3.158	0.51	0.61
WORLD	-0.293	0.215	-1.36	0.18
CMDTY	-0.034	0.081	-0.42	0.71
DVIX	0.001	0.033	0.03	0.97

5.2 EXCHANGE RATE EFFECT DURING OPEN POSITIONS

To forecast ex-ante with another procedure (which reflects possible reason for profits) the same observations as in the trading rule and then compare the results is another testing technique. Because the trading rule test has provided in US dollars we first calculate the trading rule returns in local currencies in the positions. Figure (13) shows the cumulated difference between US dollars profits and local currency profits of the trading rule returns. The difference between US dollars and local currency cumulative returns is +8.3 %. We therefore can conclude that the exchange rate effect explains 8.3 / 149.1 from returns that we gain in our positions. That is 5.6%. of the trading rule cumulative returns This not the same as the currency risk in the risk factors test, because this reflects the currency risk only during the periods we have opened positions in the MSCI country indices. That is 44 months.

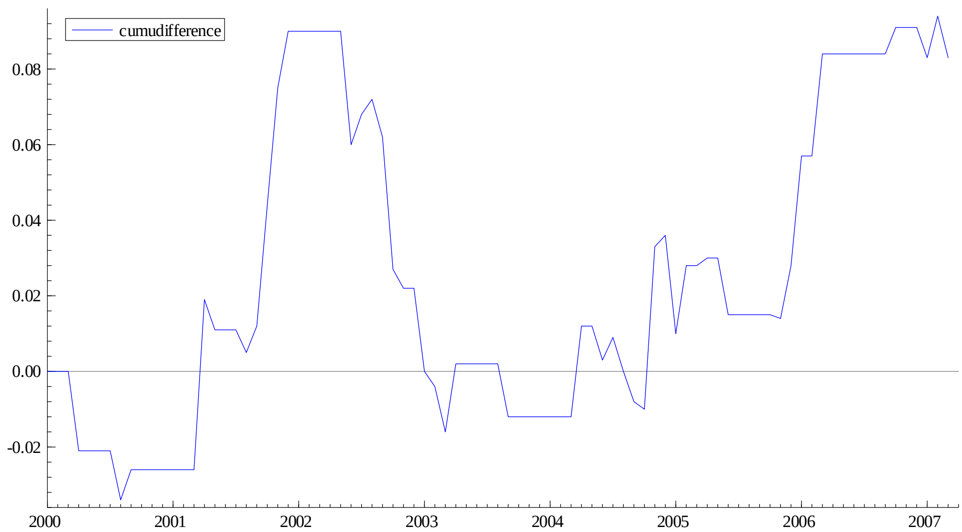

Figure 13. Cumulative difference between US dollars returns and local currency returns.

5.3 EX-ANTE TESTS: DIVIDEND YIELD, SHORT TERM RATE AND LONG TERM BOND YIELD AS PREDICTORS

Solnik (1993) argues that macroeconomic variables such as developments of dividend yield, local short-term interest rate and local long-term bond yield predict stock returns. In addition, Fama & French (1988), Campbell (1991) and Cochrane (1992) suggest that dividend yields vary over time and dividend yield variability helps to predict the risk premium of stock prices. They note that aggregate dividend yields strongly predict excess returns on a longer horizon.

We have calculated dividend yields by averaging the dividend paid over the last twelve months (data from the MSCI World index). In addition, the local short-term rate (local one-month government interest rate) and the local long-term bond yield (local ten-year government bond yield) are used as explanatory variables in this test (data from SourceOECD). For Singapore and Hong Kong these local variables were not available. For these country indices, the US short-term interest rate and the US long-term bond yield are used. Thus, we use dividend yield, local short-term rate and local long-term bond yield as a proxy for the country specific fundamentals**.** To find the forecast power of these variables the same months and indices are tested as in the original trading rule test (which has identified as forecastable in the trading rule test, total 73 positions). The model is

$$R_{t+1} = \beta_0 + \beta_1 \text{dividend yield} + \beta_2 \text{ short term rate}_t + \beta_3 \text{ bond yield}_t + e_{t+1}, \quad (5.2)$$

where R_{t+1} is the forecast of the return for period t+1 for MSCI country index for the follower country. We attempt to forecast (t+1) one-step-ahead sign of the return concerned from the dividend yield, the short term interest rate and the long term bond yield. All explanatory variables are in levels as proposed by Solnik. In addition, this test is done with OLS and with expanding windows as suggested by Solnik. That is, when new observation is available, it is included in the estimation sample. Figure (14) shows the cumulative macro factor returns with the transactions costs (same as identified in the trading rule test). The cumulative returns are -43.5 %. The hit ratio for the factor forecasts is 50.7 %. The outcome is consistent with the EMH prediction (50 %). Thus with (or without) the transaction costs the factor model cannot explains economic profits.

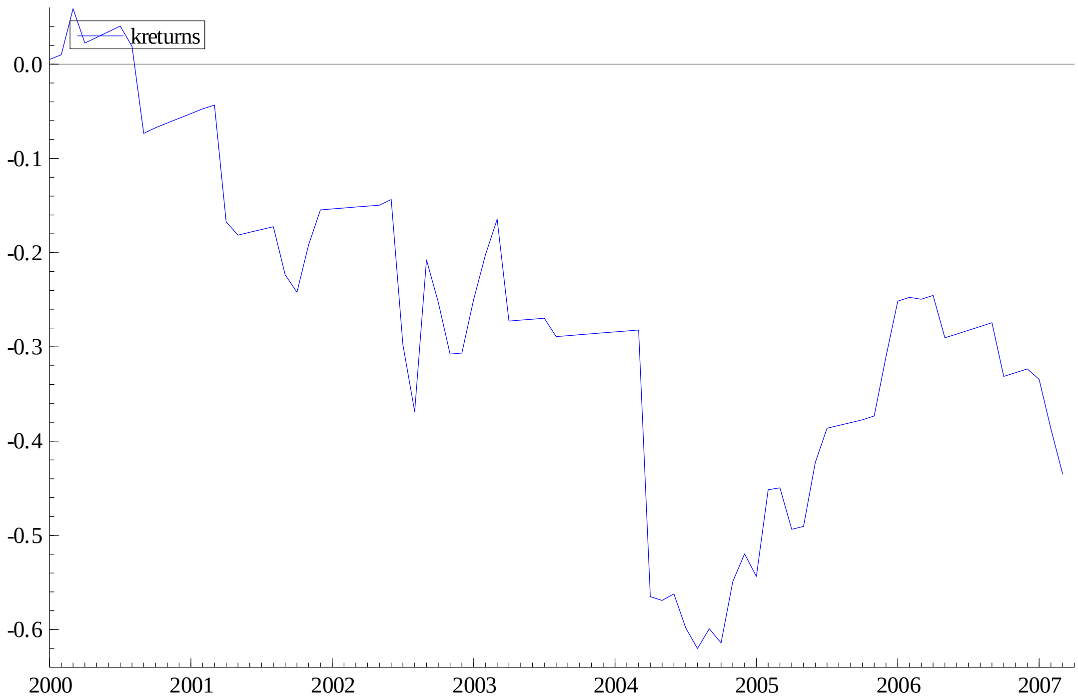

Figure 14. Cumulative macro factor returns in the same (identified) investing months as in the trading rule.

One may argue that it is more appropriate to test the factor model forecastability in the following way because these variables make rather spatial forecasts to the one-step-ahead. That is, with these variables the forecasts may not vary much in sequence. For the next test, (to keep investing space equal with other tests) we assume that MSCI USA index is a good proxy for MSCI World index. Therefore, we use US dividend yield, US short-term rate and US long-term bond yield as explanatory variables in the model (5.2) and we try to forecast the one-step-ahead sign of the return for MSCI USA index. The method is the same as in the previous test except now we invest to the index according to the model at every month (January 2000 – March 2007), instead of using the forecastable months identified by the trading rule. We take long position if the model forecasts positive return for next month and we take short position if it forecasts negative returns. The transaction costs are the same as before. That is 0.5 % per transaction and the short selling borrowing rate is 12 % in annual terms. **Now we allow multi-months positions.** This means that a position is closed only when the factor model forecasts a different sign of the return for the next investing month than the previous one(s). Thus, if it forecasts the same existing sign the long or short position is kept for the next period. In the test, the model proposes

14 long positions and 13 short positions. The longest short position is 17 months (June 2002 – October 2003) and the longest long position is 14 months (February 2006 – March 2007). Figure (14) shows the cumulative factor returns with the transaction costs. The cumulative returns are -48.5 %. Without the transaction cost returns are +9.5 %. The hit ratio for the factor forecasts is 50.8 %. The outcome is consistent with the EMH prediction (50 %). Thus, this factor model cannot explain economic profits during the testing period.

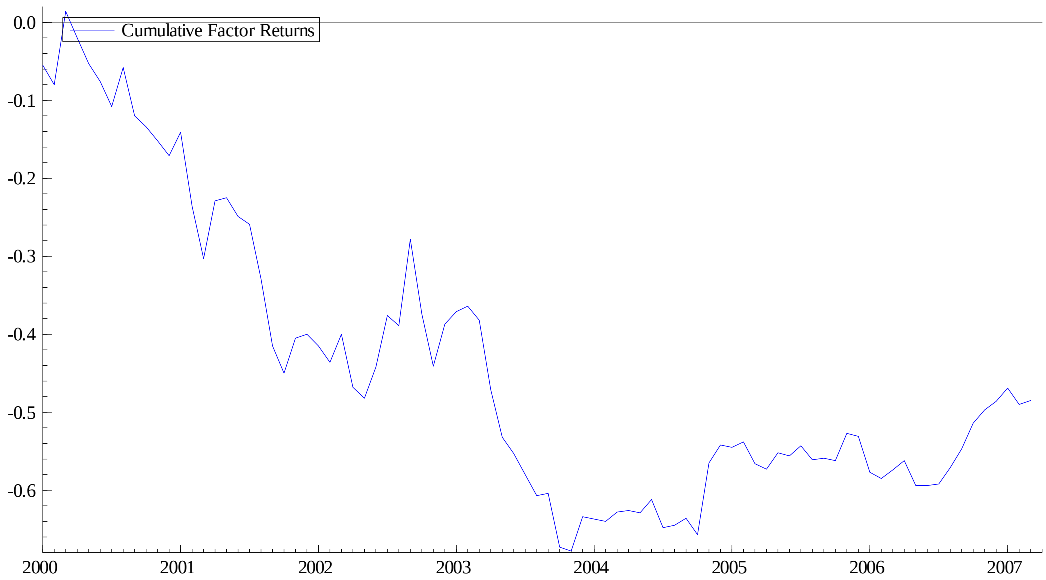

Figure 14. Cumulative US macro factor returns.

5.4 BEHAVIORAL FINANCE EXPLANATIONS FOR ECONOMIC PROFITS

Behavioral finance argues that financial phenomena can be explained using models in which some investors are not fully rational. Assets may deviate from their fundamental value. Ritter (2003) says "There is a huge psychology literature documenting that people make systematic errors in the way that they think: they are overconfident, they put too much weight on recent experience etc."

5.4.1 BEHAVIORAL THEORIES

Barberis, Shleifer & Vishny (1998), Daniel, Hirshleifer & Subrahmanyam (1998) and Hong & Stein (1999) present behavioral models based on the idea that economic profits arise because of inherent biases in the way investors interpret information. Daniel, Hirshleifer & Subrahmanyam (1998) propose that if investors are overconfident, they overreact to private information that leads to positive autocorrelation of price changes. They continue that this happens because overconfident investors also suffer from self-attribution bias. However, Daniel, Hirshleifer & Subrahmanyam argue that price moves resulting from private information are on average partially reversed in the long run. They further argue that return predictability should be stronger in firms with greater uncertainty because investors tend to be more overconfident when firms' businesses are hard to value.

Barberis, Shleifer & Vishny (1998) argue that investors think there are two regimes in the asset's earnings announcements, temporary and more permanent signs of earnings surprises. Investors assume that most of the time earnings surprises are temporary, so they underreact, for example, to positive real surprises. Barberis, Shleifer & Vishny continue that when positive earnings are followed by another positive surprise, the investors raise the likelihood that they are in the trending regime, whereas when a positive surprise is followed by a negative surprise, the investors raises the likelihood that they are in the mean-reverting regime.

Hong & Stein (1999) argue that the informed investors in their model obtain signals about future cash flows but ignore information in the past history of prices. They continue that the other investors in their model trade on the basis of a limited history of prices and, in addition, do not observe fundamental information. Hong & Stein call them momentum traders. Hong & Stein note that the information obtained by the informed investors is transmitted with a delay and hence is only partially incorporated into the prices when first revealed to the market. They argue that this part of the model contributes to underreaction, resulting in positive autocorrelation in particular prices. Therefore, it means that momentum traders can profit by trend chasing. Hong & Stein argue that eventually trend chasing leads to overreaction in the long term.

Lo (2004) proposes Adaptive Markets Hypothesis (AMH) to challenge Efficient Markets Hypothesis (EMH). He continues that (AMH) is based on an evolutionary approach to economic interactions, as well as some recent research in the cognitive neurosciences. He argues that arbitrage opportunities do exist from time to time in the

AMH. He claims that AMH implies more complex markets dynamics than EMH with cycles as well as trends, panics, manias, bubbles and crashes.

Soros (2008) claims that market prices usually express a prevailing bias rather than the correct valuation because market participants act based on imperfect understanding at all times. He continues that the illusion that markets are always right is caused by their ability to affect the fundamentals that they are supposed to reflect. Thus, Soros argues that behavioral biases may change economic fundamentals.

5.4.2 LEAD-LAG EFFECT WITHOUT ECM

The third possibility for the trading rule economic profits is that there is absolute lead-lag effect between assets concerned as proposed by Lo and McKinlay (1990). That is the cross autocorrelation puzzle. (This means that there is leader asset and follower asset in absolute prices.) Lo & McKinlay have observed that there is a lead-lag effect between large and small companies. They argue that correlations are asymmetric in the sense that returns to small companies are correlated with lagged returns on large companies, but lagged returns on small companies do not help forecasting returns on large companies. This phenomenon can be characterized as herding with a lead-lag effect. However, Lo & McKinlay do not give possible reason for the phenomenon.

To find this phenomenon the same**, (which has been identified as forecastable in the trading rule test, total 73 pairs),** lead-lag months as in the trading rule is tested with the VAR (vector autoregression) model without ECM (error correction mechanism). This means that we use only differences of logarithms of MSCI country indices. Equations (3.8 and 3.8') become then

$$\Delta X_{t+1} = \mu + \lambda_1 \Delta X_t + \kappa_1 \Delta Y_t + e_{xt+1},$$

and

$$\Delta Y_{t+1} = \mu + \lambda_2 \Delta X_t + \kappa_2 \Delta Y_t + e_{yt+1},$$

which are estimated by full information maximum likelihood method. In addition, the same rolling window size (30) is used as in the trading rule test. Thus, the only difference between the trading rule estimation and absolute lead-lag estimation is that the latter has no ECM in the model. Because a forecastable model is required to filter out white noise, the same diagnostic tests are needed with p-values ≥ 0.05, as in the trading rule ECM model. We choose the lags of variables with Akaike Information Criteria, where the model that produces the smallest AIC-value is chosen.

Figure (14) shows cumulative VAR returns with the transaction costs that is -4.3 %. We can interpret for this results that the short-run lead-lag correlation in returns alone does not help to forecast the one-step-ahead directions of changes but we need the error correction mechanism to be included to the model to make better forecasts. It therefore seems clear that in the testing period the absolute lead-lag effect cannot explain economic profits that the trading rule produces. The hit ratio for the VAR-forecasts is 50.7 %. The outcome is consistent with the EMH prediction (50 %).

Figure 14. Cumulative VAR returns in the same (identified) investing months as in the trading rule.

5.4.3 LEAD-LAG EFFECT WITH ECM IN THE MODEL

As the explanation for the remaining economic profits we propose the phenomenon where growing information uncertainty about future returns on minor investing country assets may lead to the situations where these investors take temporary role models from large financial centers for their future investing decisions, which may create a temporary herding with a lead-lag. We argue that the temporary cointegration with unidirectional Granger causality model identifies this phenomenon and it may make statistical arbitrage possible. We can interpret this because economic profits are significant and according to the estimation results (Section 5.1) risk factors (USD, BOND, CREDIT, WORLD INDEX, CMDTY & DVIX) are not source of economic profits. In addition, we have taken account foreign exchange effect in the positions and country specific fundamentals (developments of dividend yield, local short-term rate and local bond yield as a proxy). Figure (15) shows the cumulative world index + foreign exchange effect returns (+28.8 %) and the cumulative trading rule returns (+149.1 %). Their difference is +120.3 %.

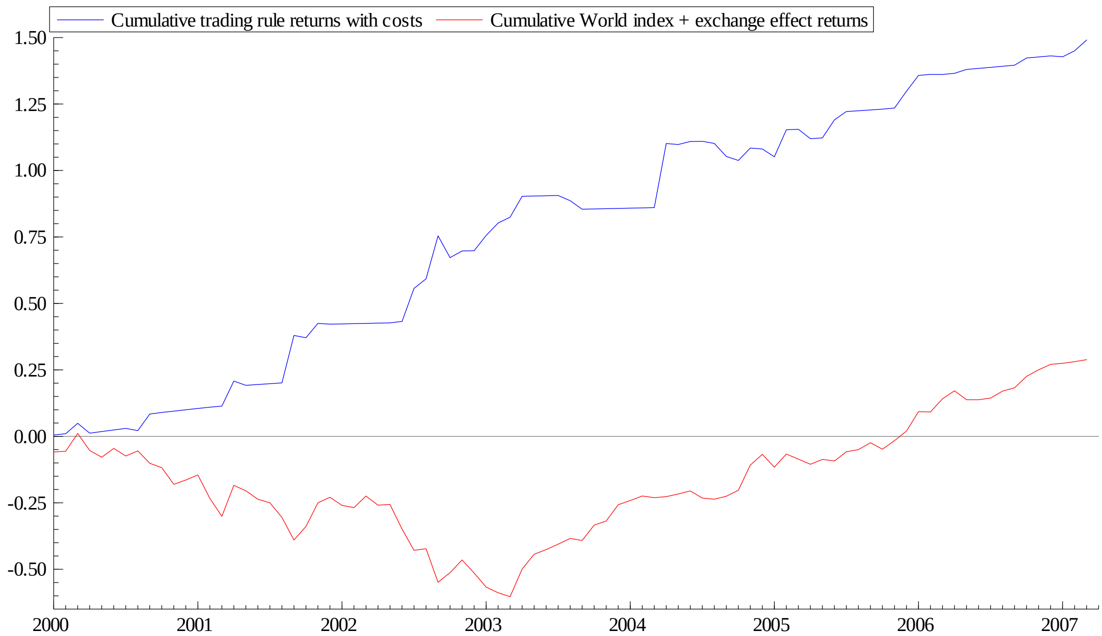

Figure 15. Cumulative trading rule returns and cumulative (world index + foreign exchange effect) returns.

6. CONCLUSIONS

According to the tests, it seems that a trading rule based on temporary cointegration with unidirectional Granger causality produces statistically significant economic profits. According to our (ex-post) risk factors test, the economic profits are not due to following risk factors: currency risk, credit risk, bond market risk, commodities, stock markets risk and volatility risk.

When we identify forecastability with the technique proposed, it means that this technique suggests that the temporary cointegration with the lead-lag effect continues to the next-step-ahead. Thus, then the probability of the continuity of unidirectional temporary cointegration to the next step is > 0.5. The technique makes a forecast about the direction of the one-step-ahead price change of the follower. In the trading rule test, the hit ratio for the forecasts is 80.8 % and it is statistically significant as p-value < 0.0001. In addition, we propose that temporary cointegration with unidirectional G-causality is consistent with herding with a lead-lag effect.

We demonstrate that in the two-asset economy where the investors have standard rational preferences but they may have biased beliefs, specific behavioral biases are sufficient conditions for existence of irrational herding with a lead-lag.

Our findings in the trading rule test are consistent with the proposition according to which the specific biased expectations may lead to a phenomenon where some investors take role models from large financial centers that lead to the temporary lead-lag effect between relative asset prices. Perfectly rational investors' behavior may also lead to return predictability such as reactions to new information from macroeconomic variables and changes of foreign exchange rates. In our (ex-ante) empirical testing, we use dividend yield, local short-term rate and local long-term bond yield as a proxy for the country specific fundamentals and these do not produce economic profits in the testing period.

We suggest that the asset prices may start to deviate from their fundamental values or from their different kinds of sentiment pricing in this particular case because of those investors' specific behavioral biases in growing uncertainty. These biases may create irrational herding behavior with a led-lag effect, which may lead to the temporary cointegration with unidirectional Granger causality. In that period, specific sentiment therefore weighs more than the fundamentals in investors' decisions. This situation lasts until the time when debiasing starts to take effect and the fundamentals or a different kind of sentiment starts to weigh again more than this specific sentiment.

However, we have to remember that an efficient market quickly learns new forecasting techniques (Timmermann & Granger 2004) when enough investors recognize new forecasting techniques. This destroys the power of a successful method. This happens because when enough investors use the method in question, its information gets into prices and it will cease to be successful.

We may have identified only specific biased expectations in the data, but our model tries to explain how investors actually behave rather than trying to create a general model of equilibrium. In addition, another common argument against behavioral models is that the results are due to data mining. We have addressed this problem by testing in real time principle.

However, in the financial markets, there is maybe difficult to separate certainly fundamental and sentiment based information that moves stock markets. In addition, private information of investors is unobservable. Therefore, the experimental finance may give further evidence of investor's biased expectations that may create unidirectional temporary cointegration. In the experimental settings, we can control the information flows that affect investors' decisions. We use this approach in our future study.

Additionally fixed window size is a rather primitive approach to optimising forecast ability, because it can be assumed that structural change is not constant over time. Furthermore, according to Pesaran and Timmermann (2002) immediately after a break, the window will tend to be too long and further away from the break the window will be too short. The problem is that no further information is used to determine possible time variation in window size. However, it seems that this problem diminishes the returns in our trading rule test and therefore strengthens our case.

REFERENCES

Alexander, Carol & Dimitriu, Anca (2005a). Indexing, cointegration and equity market regimes. *International Journal of Finance and Economics 10,* 213−231.

Alexander, Carol & Dimitriu, Anca (2005b). Indexing and statistical arbitrage: tracking error or cointegration? *Journal of Portfolio Management 31*, 50–63.

Aumann, Robert (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica 55,* 1–18.

Aumann, Robert (1998). Common priors: a reply to Gul. *Econometrica 66,* 929–938.

Balke, Nathan & Fomby, Thomas (1997). Threshold cointegration. *International Economic Review 38,* 627-645.

Banerjee, Abjihit (1992). A simple model of herd behavior. *Quarterly Journal of Economics 107*, 797−817.

Barberis, N., Shleifer, A. & Vishny R. (1998). A model of investor sentiment. *Journal of Financial Economics 49,* 307−343.

Barberis, N. Shleifer, A. & Wurgler J. (2005). Comovement. *Journal of Financial Economics 75,* 283−317.

Bernando, Antonio & Ledoit, Olivier (2000). Gain, loss and asset pricing. *Journal of Political Economy 108*, 144–172.

Black, Fischer (1986). Noise. *Journal of Finance 41*, 529−543.

Bollerslev, Tim (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics 31,* 307–327.

Bondarenko, Oleg (2003). Statistical arbitrage and securities prices. *Review of Financial Studies 16,* 875–919.

Brada, J., Kutan A. & Zhou S. (2005). Real and monetary convergence between the European Union core and recent member countries: a rolling cointegration approach. *Journal of Banking and Finance 29,* 249−270.

Campbell, John (1991). A variance decomposition of stock returns. *Economic Journal 101*, 157–179.

Chan, K.C. (1988). On the contrarian investment strategy. *Journal of Business 61,* 147–163.

Cipriani, Marco & Guarino, Antonio (2008). Herd behavior and contagion in financial markets. *B. E. Journal of Theoretical Economics 8,* 1−56.

Cochrane, John (1992). Explaining the variance of price dividend ratios. *Review of Financial Studies 5,* 243–280.

Cochrane, John & Saá-Requejo, Jesus (2000). Beyond arbitrage: “Good-Deal” asset pricing bounds in incomplete markets. *Journal of Political Economy 108,* 79–119.

Conrad, Jennifer & Kaul, Gautam (1998). An anatomy of trading strategies. *Review of Financial Studies 5,* 243–280.

Daniel, K., Hirshleifer, D. & Subrahmanyam A. (1998). Investor psychology and security market under- and overreaction. *Journal of Finance 53,* 1839−1885.

Dasgupta, Amil & Prat, Andrea (2008). Information aggregation in financial markets with career concerns. *Journal of Economic Theory 143*, 83−113.

De Bondt, William. & Thaler, Richard. (1985). Does the stock market overreact? *Journal of Finance 35*, 793−805.

Dunis, Christian & Ho, Richard (2005). Cointegration portfolios of European equities for index tracking and market neutral strategies. *Journal of Asset Management 6*, 33–52.

Edwards, William (1968). Conservatism in human information processing, in *B. Kleimutz ed. Representation of Human Judgement, John Wiley & Sons.*

Engle, Robert & Granger, Clive (1987). Co-integration and error-correction: Representation, estimation and testing. *Econometrica 55*, 251−276.

Fair, Ray & Shiller, Robert (1990). Comparing information in forecasts from econometric models. *American Econometric Review 80,* 375−389.

Fama, Eugene & French, Kenneth (1988a). Permanent and temporary components of stock prices. *Journal of Political Economy 96*, 246–273.

Fama, Eugene & French, Kenneth (1988b). Dividend yields and expected stock returns. *Journal of Financial Economics* 22, 3–26.

Fama, Eugene & French, Kenneth (1992). The cross-section of expected returns. *Journal of Finance 47*, 427–465.

Fama, Eugene & French, Kenneth (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*, 3–56.

Fung, William & Hsieh, David (1997). Empirical characteristics of dynamic trading strategies: The case of hedge funds. *Review of Financial Studies 10,* 275–302.

Gatev, E., Goetzmann, W. & Rouwenhorst, G. (2006). Pairs trading: performance of a relative value arbitrage rule. *Review of Financial Studies 19,* 797–827.

Getmansky, M., Lo A. & Makarov, I. (2004). An econometric analysis of serial correlation and illiquidity in hedge-funds returns. *Journal of Financial Economics 74,* 529−609.

Granger, Clive (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37,* 424−438.

Granger, Clive (1980). Testing for causality. *Journal of Economic Dynamics and Control 2*, 329−352.

Granger, Clive (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics 48*, 213–228.

Granger, Clive (1988). Some recent developments in a concept of causality. *Journal of Econometrics 39,* 199−211.

Granger, Clive (2003). Some aspects of causal relationships. *Journal of Econometrics 112,* 69−71.

Granger, Clive & Siklos, Pierre (1997). Regime-Sensitive cointegration with an application to interest-rate parity. *Macroeconomic Dynamics 1,* 640–657.

Granger, Clive & Timmermann, Allan (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting 20*, 15−27.

Hansen, Bruce & Seo, Byeongseon (2002). Testing for two regime threshold cointegration in vector error correction models. *Journal of Econometrics 110.* 293–318.

Harrison, James & Kreps, David (1979). Martingales and arbitrage in multiperiod securities prices. *Journal of Economic Theory 20,* 381–408.

Hirshleifer, David & Teoh, Siew Hong (2003). Herd behavior and cascading in capital markets: a review and synthesis. *European Financial Management 9,* 25−66.

Hong, Harrison & Stein, Jeremy (1999). A unified theory of under reaction, momentum trading and overreaction in asset markets. *Journal of Finance 54,* 2143−2184.

Jegadeesh, Narasimhan (1990). Evidence of predictable behaviour of security returns. *Journal of Finance 45,* 881–898.

Jegadeesh, Narasimhan & Titman, Sheridan (1993). Returns of buying winners and selling losers: implications of stock market efficiency. *Journal of Finance 54,* 2143–2184.

Jegadeesh, Narasimhan & Titman, Sheridan (1995). Overreaction, Delayed reaction and contrarian profits. *Review of Financial Studies 8,* 973−993.

Jegadeesh, Narasimhan & Titman, Sheridan (2001). Profitability of momentum strategies: An evolution of alternative explanations. *Journal of Finance 56,* 699−720.

Jensen, Michael (1969). The performance of mutual funds in the period 1955−1964. *Journal of Finance 23*, 389-416.

Johansen, Soren (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control 12*, 231–254.

Johansen, Soren (1991). Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. *Econometrica 59,* 1551–1580.

Johansen, Soren & Juselius, Katarina (1990). Maximum likelihood estimation and inference on cointegration, with application for the demand for money. *Oxford Bulletin of Economics and Statistics 52,* 169–210.

J.P. Morgan (1996). *Riskmetrics technical document,* 4^{th} ed. J.P. Morgan, New York.

Kahneman, Daniel. & Tversky, Amos. (1974). Judgement under uncertainty: Heuristics and biases. *Science 185,* 1124–1131.

Kasa, Kenneth (1992). Common stochastic trends in international stock markets. *Journal of Monetary Economics 29,* 95 – 124.

Lehmann, Bruce (1990). Fads, martingales and market efficiency. *Quarterly Journal of Economics 50*, 1–28.

Lo, Andrew (2004). The Adaptive market hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management 30*, 15−29.

Lo, Andrew (2008). *Hedge funds, an analytic perspective*. Princeton University Press, New Jersey.

Lo, Andrew & McKinlay, Craig (1988). Stock market prices do not follow random walk: evidence from a simple specification test. *Review of Financial Studies 1,* 41–66.

Lo, Andrew & McKinlay, Craig (1990). When are contrarian profits due stock market overreaction? *Review of Financial Studies 3,* 175−208.

Lo, Ming & Zivot, Eric (2001). Threshold cointegration and nonlinear adjustment to the law of one price. *Macroeconomic Dynamics 5,* 533–576.

Malkiel, Burton (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives 17,* 59–82.

Murdock, Bennett (1962). The serial position effect on free retrial. *Journal of Experimental Psychology 64,* 482–488.

Nelson, Daniel (1990). Stationary and persistence in the GARCH (1,1) model. *Econometric Theory 6*, 318–334.

Oskamp, Stuart (1965). Overconfidence in case study judgements. *Journal of Consulting Psychology*. 261−265.

Pesaran, Hashem & Timmermann, Allan (1992). A simple non-parametric test of predictive performance. *Journal of Business and Economic Statistics 10,* 461–465.

Pesaran, Hashem & Timmermann, Allan (1995). Predictability of stock returns: robustness and economic significance. *Journal of finance 50,* 1201−1228.

Pesaran, Hashem & Timmermann, Allan (2002). Market timing and return prediction under model instability. *Journal of Empirical Finance 9,* 495−510.

Rangvid, Jesper & Sorensen Carsten (2002). Converge in the ERM and declining numbers of common stochastic trends. *Journal of Emerging Market Finance 1,* 183−213.

Richards, Anthony (1995). Comovements in national stock market returns: evidence of predictability, but not cointegration. *Journal of Monetary Economics 36,* 631–654.

Ritter, Jay (2003). Behavioral Finance. *Pacific-Basin Finance Journal 11,* 429−437.

Saikkonen, Pentti (2008). Stability of regime switching error correction models under linear cointegration. *Econometric Theory 24,* 294–318.

Samuelson, Paul (1965). Proof that properly anticipated prices fluctuate randomly, *Industrial Management Review 6*, 4–12.

Savage, Leonard. (1964). *The Foundation of Statistics.* John Wiley & Sons, New York.

Scharfstein, David & Stein, Jeremy (1990). Herd Behavior and Investment. *American Economic Review 80,* 465–479.

Sharpe, William (1992). Asset allocation: Management style and performance measurement. *Journal of Portfolio Management Winter,* 7–19.

Shefrin, Hersh (2000). *Beyond Greed and Fear*. Oxford University Press, New York.

Shewhart, William (1939). *Statistical method from the viewpoint of quality control*, Dover, NY.

Shiller, Robert (2000). *Irrational exuberance.* Princeton University Press, New York.

Shleifer, Andrei (2000). *Inefficient Markets.* Oxford University Press, New York.

Shleifer, Andrei & Vishny, Robert (1997). The Limits of arbitrage. *Journal of Finance 51,* 35–55.

Solnik, Bruno (1993). The performance of international asset allocation strategies using conditioning information. *Journal of Empirical Finance 1*, 33-55.

Soros, George (2008). *The new paradigm for financial markets.* Public Affairs. New York.

Subrahmanyam, Avanidhar (2005). On distinguishing between rationales for short horizon predictability of stock returns. *Financial Review 40*, 11–35.

Subrahmanyam, Avanidhar (2008). Behavioural finance: A Review and synthesis. *European Financial Management 14,* 12–29.

Summers, Lawrence (1986). Does the stock market rationally reflect fundamental values? *Journal of Finance 41,* 591–601.

Swanson, Norman (1998). Money and output viewed through a rolling window. *Journal of Monetary Economics 41,* 455−473.

Swanson, Norman & Zeng, Tian (1998). Predictive evaluation of econometric forecasting models in commodity future markets. *Studies in Nonlinear Dynamics and Econometrics 2*, 159−177.

Tsay, Ruey (1998). Testing and modelling multivariate threshold models. *Journal of the American Statistical Association 93*, 1188–1202.

Zhang, Frank (2006). Information uncertainty and stock returns. *Journal of Finance 61*, 105–138.

WORKING PAPERS

Alexander, Carol (1999). Optimal hedging using cointegration. *Philosophical Transactions of the Royal Society, Series A 357,* 2039–2058.

Heracleous, M., Koutris A. & Spanos A. (2005). Testing for structural change: a misspecification testing perspective.

Granger, Clive & Yoon, Gawon (2002). Hidden cointegration. *University of California, San Diego Department of Economics working paper 2002-2.*

Pesaran, Hashem & Timmermann, Allan (2006). Testing dependence among serially correlated multi-category variables. *CESifo Working Paper No. 177*

APPENDIX I

Johansen (1988 & 1991) and Johansen & Juselius (1990) propose that cointegration rank r and linear restrictions on the cointegrating relations β' and on the adjustment coefficients α can be estimated using the maximum likelihood method. Thus we estimate rank r and unique relations of $\Pi = \alpha\beta'$. First, we construct vector autoregression (VAR): $y_t = \Gamma_1 y_{t-1} + \Gamma_2 y_{t-2} + \ldots \Gamma_p y_{t-p} + e_t$. Now let z_t denote the vector of n(p – 1) variables, where n is number of variables in the VAR model. Thus $z_t = [\Delta y_{t-1}, \Delta y_{t-2} + \ldots . \Delta y_{t-p+1}]$. Next we define that 0t is regression of Δy_t on z_t and 1t is regression of y_{t-p} on z_t. This yields the residuals $\mathbf{R}_{0t}$ and $\mathbf{R}_{1t}$. Next we compute the second moments of all these residuals in a sample, denoted $\mathbf{S_{00}}$, $\mathbf{S_{01}}$ and $\mathbf{S_{11}}$ where:

$$S_{ij} = 1/T \sum_{t=1}^{T} R_{it} R_{jt}' \text{ for } i, j = 0,1.$$

Next we solve $|\boldsymbol{\lambda} \mathbf{S_{11}} - \mathbf{S_{10}} \mathbf{S_{00}}^{-1} \mathbf{S_{01}}| = \mathbf{0}$ for the r largest eigenvalues $\mathbf{1} > \boldsymbol{\lambda}^{\wedge}_{\mathbf{1}} > \ldots . > \boldsymbol{\lambda}^{\wedge}_{\mathbf{r}} \ldots > \boldsymbol{\lambda}^{\wedge}_{\mathbf{n}} > \mathbf{0}$ and the corresponding eigenvectors:

$\boldsymbol{\beta}^{\wedge} = (\boldsymbol{\beta}^{\wedge}_{\mathbf{1}}, \ldots .., \boldsymbol{\beta}^{\wedge}_{\mathbf{r}})$ normalized by $\boldsymbol{\beta}^{\wedge\prime} \mathbf{S_{11}} \boldsymbol{\beta}^{\wedge} = \mathbf{1_r}$.

Then, tests of hypothesis of r cointegrating vectors can be based on the trace statistic:

$$\eta_r = -T \sum_{i=r+1}^{n} \log (1 - \lambda^{\wedge}_i).$$

To estimate unique cointegration relations, (for example, which includes weak exogeneity in the cointegrated system) we test restrictions (of $H_0: \Pi = \alpha\beta'$ versus $H_1: \Pi \rightarrow \beta = \varphi H$ *and* $\alpha = \psi A$ where H is (s * r) and A is (m * r) where $r \leq s \leq n$ and $r \leq m \leq n$) with the likelihood ratio test. The likelihood ratio test statistic is given by

$$-2 \log (Q; H_1 \mid H_0) = T \sum_{i=1}^{r} \log \{ (1 - \lambda^{\wedge}_i (H_1)) / (1 - \lambda^{\wedge}_i (H_0))\}.$$

More information of the procedure see Johansen (1988 & 1991) and Johansen & Juselius (1990).

APPENDIX II

Potential lead lag –pairs (81) that we tested during period 1.1.2000-31.3.2007.

USA→ Finland
USA→ Norway
USA→ Sweden
USA→ Denmark
USA→ Austria
USA→ Belgium
USA→ Italy
USA→ Spain
USA→ Greece
USA→ Ireland
USA→ Portugal
USA→ France
USA→ Holland
USA→ Singapore
USA→ Australia
USA→ Canada
USA→ Hong Kong
USA→ New Zealand
UK→ Finland
UK→ Norway
UK→ Sweden
UK→ Denmark
UK→ Austria
UK→ Belgium
UK→ Italy
UK→ Spain
UK→ Greece
UK→ Ireland
UK→ Portugal
UK→ France
UK→ Holland
UK→ Singapore
UK→ Australia
UK→ Canada
UK→ Hong Kong
UK→ New Zealand
Switzerland→ Finland
Switzerland→ Norway
Switzerland→ Sweden
Switzerland→ Denmark
Switzerland→ Austria
Switzerland→ Belgium
Switzerland→ Italy
Switzerland→ Spain
Switzerland→ Greece
Switzerland→ Ireland
Switzerland→ Portugal
Switzerland→ France
Switzerland→ Holland
Switzerland→ Singapore
Switzerland→ Australia
Switzerland→ Canada
Switzerland→ Hong Kong
Switzerland→ New Zealand
Germany→ Finland
Germany→ Norway
Germany→ Sweden
Germany→ Denmark
Germany→ Austria
Germany→ Belgium
Germany→ Italy
Germany→ Spain
Germany→ Greece
Germany→ Ireland
Germany→ Portugal
Germany→ France
Germany→ Holland
Japan→ Singapore
Japan→ Australia
Japan→ Australia
Japan→ New Zealand
Hong Kong→ Canada
World index→ USA
World index→ UK
World index→ Japan
World index→ Germany
USA→ UK
USA→ Switzerland
USA→ Japan
USA→ Germany

APPENDIX III

Hit ratio data

	Forecasted	Observed	
1.	UP	UP	HIT
2.	UP	DOWN	FALSE
3.	DOWN	DOWN	HIT
4.	DOWN	DOWN	HIT
5.	UP	UP	HIT
6.	UP	UP	HIT
7.	DOWN	DOWN	HIT
8.	DOWN	DOWN	HIT
9.	UP	UP	HIT
10.	UP	UP	HIT
11.	DOWN	UP	FALSE
12.	DOWN	DOWN	HIT
13.	DOWN	DOWN	HIT
14.	DOWN	DOWN	HIT
15.	DOWN	DOWN	HIT
16.	UP	UP	HIT
17.	DOWN	DOWN	HIT
18.	DOWN	DOWN	HIT
19.	DOWN	UP	FALSE
20.	DOWN	UP	FALSE
21.	UP	UP	HIT
22.	UP	UP	HIT
23.	DOWN	DOWN	HIT
24.	DOWN	DOWN	HIT
25.	DOWN	DOWN	HIT
26.	DOWN	DOWN	HIT
27.	DOWN	UP	FALSE
28.	DOWN	DOWN	HIT
29.	UP	UP	HIT
30.	UP	UP	HIT
31.	DOWN	DOWN	HIT
32.	DOWN	UP	FALSE
33.	DOWN	DOWN	HIT
34.	DOWN	DOWN	HIT
35.	DOWN	DOWN	HIT
36.	UP	UP	HIT
37.	UP	UP	HIT
38.	UP	UP	HIT
39.	DOWN	DOWN	HIT
40.	DOWN	DOWN	HIT
41.	DOWN	UP	FALSE
42.	UP	UP	HIT
43.	DOWN	DOWN	HIT

44.	UP	UP	HIT
45.	UP	UP	HIT
46.	DOWN	UP	FALSE
47.	UP	UP	HIT
48.	UP	UP	HIT
49.	UP	UP	HIT
50.	DOWN	UP	FALSE
51.	UP	DOWN	FALSE
52.	UP	DOWN	FALSE
53.	UP	UP	HIT
54.	UP	DOWN	FALSE
55.	UP	DOWN	FALSE
56.	UP	DOWN	FALSE
57.	UP	UP	HIT
58.	UP	UP	HIT
59.	UP	UP	HIT
60.	UP	UP	HIT
61.	UP	UP	HIT
62.	UP	UP	HIT
63.	UP	UP	HIT
64.	UP	UP	HIT
65.	UP	UP	HIT
66.	DOWN	DOWN	HIT
67.	UP	UP	HIT
68.	UP	UP	HIT
69.	UP	UP	HIT
70.	DOWN	DOWN	HIT
71.	UP	UP	HIT
72.	UP	UP	HIT
73.	UP	UP	HIT

HIT RATIO SUMMARY

forecasted	**HIT**	**FALSE**	total
UP	35	6	41
DOWN	24	8	32
total	59	14	**73**