MARTINGALE PRICING THEORY

A Thesis Presented

by

Nathaline Hajj Moussa

to

The Faculty of Natural and Applied Sciences

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Mathematics

Notre Dame University-Louaize

Zouk Mosbeh, Lebanon

May 2019

Notre Dame University-Louaize,
Zouk Mosbeh, Lebanon

Department of Mathematics and Statistics

Nathaline Hajj Moussa

We, the thesis committee for the above candidate for the Master of Science degree, hereby recommend acceptance of this thesis.

John Haddad – Thesis Advisor
Associate Professor, Department of Mathematics and Statistics

Bassem Ghalayini – First Reader
Associate Professor, Department of Mathematics and Statistics

Holem Saliba – Second Reader
Associate Professor, Department of Mathematics and Statistics

This thesis is accepted by the Faculty of Natural and Applied Sciences.

Professor George Eid
Dean of the Faculty of Natural and
Applied Sciences

Abstract of the Thesis

MARTINGALE PRICING THEORY

by

Nathaline Hajj Moussa

Master of Science

in

Mathematics

Notre Dame University-Louaize,

Zouk Mosbeh,
Lebanon

2019

From ages to ages there had been expectation of individuals on a specific predictions and future occurrences. So also in a game, different participant that involves in those specified game have their various expectations of the results or the output of the game they are involved in. That is why we need a mathematical theory that helps in prediction of the future expectations in our day to day activities. Therefore the Martingale Theory is a very good theory that explains and dissects the expectation of a gamer in a given

game of chance. So in this thesis, we shall talk about the Martingale Theory expressing the expectations of a gamer in a game of chance, and also discuss the gaming strategies so as to enlighten everyone involved in a specific game their required expectation after proper understanding of the Martingale Theory. Also this thesis examines testing Martingale Difference Hypothesis (MDH) and related statistical inference issues and it discusses the developments of tests and applies them to exchange rate data.

To my family.

Contents

Acknowledgements

I would like to express my gratitude to my supervisor Dr.John Haddad for his advisement and support throughout all the thesis period. Without his guidance, it would be impossible for me to overcome all the difficulties. Furthermore, I appreciate highly his great efforts in my literature review, research, and the thesis writing. Also,my sincere gratitude to all professors of Master program Financial Mathematics for their effort in providing us with high quality knowledge.

Last but not least,I wish to thank my family for their everlasting support.

Chapter 1

Martingale(probability theory)

1.1 Introduction

Martingale is a betting strategy that was traced back as at 18 centuries in France. This strategy was introduced for a game in which a specific gambler wins his stake if a coin comes up heads and loses it if same coin comes up tails. The gambler needs to double his bet after every loss since he/she is not ready to loose nor give up and his/her aim is to recover all previous losses plus win and gain a profit that is equivalent to the original stake. This same Martingale strategy has been applied to some other games like the roulette, as the probability of hitting either red or black is close to 0.5.

We can also describe a Martingale as a model of a fair game in which the knowledge of the past events or the knowledge of the already known result of the game can never help to predict the result or the mean of the expected winnings. Consequently, a Martingale is a sequence of random variables or rather a stochastic process for which, at a given time in the realized sequence, the expectation of the next value in the sequence is equal to the present observed value inconsequential of the knowledge of all previously observed values.

On the other hand, in a non-martingales process, we may still have a situation where the expected value of the process at one time is equal to the expected value of the process at the next time. However, knowledge of the previous outcomes, for instance, "the previous cards drawn out from a set of cards" may be able to reduce the uncertainty of upcoming outcomes. Thus, the expected or resulting value of the next outcome given a definite knowledge of the present and all previous outcomes may definitely be higher than the current outcome provided we use the said winning strategy. Martingale does not include the possibility of the winning strategies based on already known game history, and thus making the system a model of fair games. Since a gambler with inexhaustible measure of wealth will almost surely flip head, with this said reason, the Martingale betting strategy was concluded to be as a sure system of gaming by those who recommended it. Even though none of the gamblers possesses an inexhaustible wealth, and the exponential movement of the bets placed would eventually bankrupt the gambler and the gamblers who chose to use the Martingale system often wins a minute net reward, thus appearing to have a faultless and accurate strategy. However, the gambler's expected results and values mostly ends up being zero (or even less than zero) because the small probability that he will suffer an unimaginable loss exactly measures up and balance up with his gain. (In a casino, the expected result of a gambler is negative, simply because to the house's edge.) The possibility of catastrophic loss may not really be small since the bet size always rises in an exponential rate. The fact that strings of consecutive losses definitely occur more often than just an ordinary intuitional suggestions, can make the gambler go bankrupt quickly.

1.2 Martingale Definition

1.2.1 Definition

A stochastic process $\{X_n; n = 0, 1, ...\}$ for n=0,1,... is a martingale if

1. $E[|X_n|] < \infty$

2. $E[X_{n+1}|X_0...X_n] = X_n$

Also we say $\{X_n\}$ is a martinagle with respect to $\{Y_n\}$ if for n=0,1,... $\{X_n; n = 0, 1, ...\}$ and $\{Y_n; n = 0, 1, ...\}$ are stochastic processes then

1. $E[|X_n] < \infty$

2. $E[X_{n+1}|Y_0, ..., Y_n] = X_n$

Or we can define it for the continous case like the following:

1. $E[|X_n|] < \infty$

2. $E[X_t|\{X_0, ..X_s\}] = X_s, \forall 0 \leq s \leq t$

1.2.2 Examples

(a) A gambler's fortune (capital) is a martingale if all the betting games which the gambler plays are fair. To be more specific: suppose X_n is a gambler's fortune after n tosses of a fair coin, where the gambler wins $1 if the coin comes up heads and loses $1 if it's tails. The gambler's conditional expected fortune after the next trial, given the history, is equal to his present fortune. This sequence is thus a martingale.

(b) Sums of Independent Random Variables:

Let $Y_0 = 0$ and $Y_1, Y_2, ...$ be independent random variables with $E[|Y_n| = 0]$ for all n. If $X_0 = 0$ and $X_n = Y_1 + ...Y_n$ for $n > 1$, then $\{X_n\}$ is a martingale with respect to $\{Y_n\}$.

- $E[|X_n|] < E[|Y_1|] + ...E[|Y_n|] < \infty$
- $E[X_{n+1}|Y_0, ..., Y_n] = E[X_n + Y_{n+1}|Y_0, ..., Y_n]$

  $= E[X_n|Y_0, ..., Y_n] + E[Y_{n+1}|Y_0, ..., Y_n \;\; = X_n + E[Y_{n+1}] = X_n$

  (because if the independence assumption on $\{Y_i\}$

(c) The variance of sums as a martingale:

Let $Y_0 = 0$ and $Y_1, ..., Y_n$ be iid random variables with$E[Y_k] = 0$ and $E[Y_K^2] = \sigma^2$, $k = 1, 2, ...$ and let $X_0 = 0$ and $X_n = (\sum_{k=1}^{n} Y_k)^2 - n\sigma^2$ then $E[|X_n|] < 2n\sigma^2 < \infty$, and

$$E[X_{n+1}|Y_0, ..., Y_n]$$

$$= E[(Y_{n+1} + \sum_{k=1}^{n} Y_k)^2 - (n+1)\sigma^2|Y_0, ..., Y_n]$$

$$= E[Y_{n+1}^2 + 2Y_{n+1}\sum_{k=1}^{n} Y_k + (\sum_{k=1}^{n} Y_k)^2 - n\sigma^2 - \sigma^2|Y_0, ..., Y_n]$$

$$= E[Y_{n+1}^2|Y_0, ..., Y_n] + 2(\sum_{k=1}^{n} Y_k)E[Y_{n+1}|Y_0, ..., Y_n] + X_n + n\sigma^2 - n\sigma^2 - \sigma^2]$$

$$= X_n$$

(d) Doob's martingale process:

A Doob martingale (also known as a Levy martingale) is a mathematical construction of a stochastic process which approximates a given random variable and has the martingale property with respect to the given filtration. It may be thought of as the evolving sequence of best approximations to the random variable based on information accumulated up to a certain time. Let$Y_0, Y_1, ...$ be an arbitrary sequence of random variables and suppose X is a random variable satisfying $E[|X|] < \infty$. Then

$$X_n = E[X|Y_0, ..., Y_n]$$

forms a martingale with respect to $\{Y_n\}$, called Doob's process.

First $E[|X_n|] = E\{|E[X|Y_0, ..., Y_n]|\} \leq E\{E[|X|Y_0, .., Y_n]\} = E[|X|] < \infty$

Second and last, by the law of total probability for conditional expectations

$E[X_{n+1}|Y_0, ..., Y_n]$

$= E\{E[X|Y_0, ..., Y_{n+1}|Y_0, ..., Y_n\}$

$= E[X|Y_0, ..., Y_n] = X_n$

(e) De Moivre's martingale:

Now suppose the coin is unfair, i.e., biased, with probability p of coming up heads and probability $q = 1 - p$ of tails. Let

$$X_{n+1} = X_n + 1$$

in case of heads or

$$X_{n+1} = X_n - 1$$

with + in case of tails. Let

$$Y_n = (q/p)^{X_n}$$

then $\{Y_n : n = 1, 2, ...\}$ is a martingale with respect to $\{X_n : n = 1, 2, ...\}$. To show this

$$E[Y_{n+1}|X_1, ..., X_n]$$

$$= p(q/p)^{X_{n+1}} + q(q/p)^{X_{n-1}}$$

$$= p(q/p)(q/p)^{X_n} + q(p/q)(q/p)^{X_n}$$

$$= q(q/p)^{X_n} + p(q/p)^{X_n}$$

$$= (p + q)(q/p)^{X_n}$$

$$= (q/p)^{X_n} = Y_n$$

(f) Polya's urn:

contains a number of different coloured marbles; at each iteration a marble is randomly selected from the urn and replaced with several more of that same colour. For any given colour, the fraction of marbles in the urn with that colour is a martingale. For example, if currently 95 % of the marbles are red then, though the next iteration is more likely to add red marbles than another color, this bias is exactly balanced out by the fact that adding more red marbles alters the fraction much less significantly than adding the same number of non-red marbles would.

(g) Likelihood ratio testing:

A random variable X_n is thought to be distributed according either to probability density f or to a different probability density g. A random sample $X_1, ..., X_n$ is taken. Let Y_n be the "likelihood ratio"

$$Y_n = \prod_{i=1}^{n} \frac{g(X_i)}{f(X_i)}$$

1.3 Supermartingales and submartingales

There are two popular generalizations of a martingale that also include cases when the current observation X_n is not necessarily equal to the future conditional expectation$E[X_{n+1}|X_1, ..., X_n]$ but instead an upper or lower bound on the conditional expectation. These definitions reflect a relationship between martingale theory and potential theory, which is the study of harmonic functions. Just as a continuous-time martingale satisfies $E[X_t|\{X_T : T \leq S\}] - X_S = 0, \forall\ s \leq t$, a harmonic function f satisfies the partial differential equationΔ f = 0 where Δ is the Laplacian operator. Given a Brownian motion process W_t and a harmonic function f, the resulting process $f(W_t)$ is also a martingale.

(a) A discrete-time submartingale is a sequence $X_1, X_2, ...$ of integrable random variables satisfying

$$E[X_{n+1}|X_1, ..., X_n] \geq X_n$$

Likewise, a continuous-time submartingale satisfies

$$E[X_t|\{X_T : T \leq s\}] \geq X_s \forall s \leq t$$

In potential theory, a subharmonic function f satisfies $\Delta f \geq 0$. Any subharmonic function that is bounded above by a harmonic function for all points on the boundary of a ball are bounded above by the harmonic function for all points inside the ball. Similarly, if a submartingale and a martingale have equivalent expectations for a given time, the history of the submartingale tends to be bounded above by the history of the martingale. Roughly speaking, the prefix "sub-" is consistent because the current observation X_n is less than (or equal to) the conditional expectation $E[X_{n+1}|X_1..X_n]$. Consequently, the current observation provides support from below the future conditional expectation, and the process tends to increase in future time.

(b) Analogously, a discrete-time supermartingale satisfies

$$E[X_{n+1}|X_1, .., X_n] \leq X_n$$

Likewise, a continuous-time supermartingale satisfies

$$E[X_t|\{X_T : T \leq s\}] \leq X_s, \forall s \leq t$$

In potential theory, a superharmonic function f satisfies $\Delta f \leq 0$. Any superharmonic function that is bounded below by a harmonic function for all points on the boundary of a ball are bounded below by the harmonic function for all points inside the ball. Similarly, if a supermartingale and a martingale have equivalent expectations for a given time, the history of the supermartingale tends to be bounded below by the history of the martingale. Roughly speaking, the prefix "super-" is consistent because the current observation X_n is greater than (or equal to) the conditional expectation $E[X_{n+1}|X_1, ..., X_n]$. Consequently, the current observation provides support from above the future conditional expectation, and the process tends to decrease in future time.

1.3.1 Examples

i. Every martingale is also a submartingale and a supermartingale. Conversely, any stochastic process that is both a submartingale and a supermartingale is a martingale.

ii. Consider again the gambler who wins $1 when a coin comes up heads and loses $1 when the coin comes up tails. Suppose now that the coin may be biased, so that it comes up heads with probability p.

- If p is equal to 1/2, the gambler on average neither wins nor loses money, and the gambler's fortune over time is a martingale.
- If p is less than 1/2, the gambler loses money on average, and the gambler's fortune over time is a supermartingale.
- If p is greater than 1/2, the gambler wins money on average, and the gambler's fortune over time is a submartingale.

iii. A convex function of a martingale is a submartingale, by Jensen's inequality. For example, the square of the gambler's fortune in the fair coin game is a submartingale (which also follows from the fact that $X_n^2 - n$ is a martingale). Similarly, a concave function of a martingale is a supermartingale.

1.4 Stopping time and the optional stoppig theorem

A stopping time with respect to a sequence of random variables $X_1, X_2, ...$ is a random variable T with the property that for each t, the occurrence or non-occurrence of the event T= t depends only on the values of $X_1, X_2, ...X_t$. The intuition behind the definition is that at any particular time t, you can look at the sequence so far and tell if it is time to stop. An example in real life might be the time at which a gambler leaves the gambling table, which might be a function of his previous winnings (for example, he might leave only when he goes broke), but he can't choose to go or stay based on the outcome of games that haven't been played yet. One of the basic properties of martingales is that, if $(X_t)_{t>0}$ is a (sub-/super-) martingale and T is a stopping time, then the corresponding stopped process $(X_t^T)_{t>0}$ defined by $X_t^T = X_{min\{T,t\}}$ is also a (sub-/super-) martingale.

The concept of a stopped martingale leads to a series of important theorems, including, for example, the optional stopping theorem (or Doob's optional sampling theorem) which states that, under certain conditions, the expected value of a martingale at a stopping time is equal to its initial value; since martingales can be used to model the wealth of a gambler participating in a fair game, the optional stopping theorem says that, on average, nothing can be gained by stopping play based on the information obtainable so far (i.e., without looking into the future). Certain conditions are necessary for this result to hold true. In particular, the theorem applies to doubling strategies.In other words; Let $(M_n)_{n\geq 0}$ be a martingale and let T be a stopping time. Suppose that at least one the following conditions hold:

1. $T \leq n$ for some n

2. $T < \infty$ and $|M_n| \leq C$ whenever $n \leq T$

Then $E[M_T] = E[M_0]$

The optional stopping theorem is an important tool of mathematical finance in the context of the fundamental theorem of asset pricing.

1.5 Martingale convergence theorem

Martingale convergence theorem is a special type of theorem, since the convergence follows from structural properties of the sequence of random variables.

- Let $\{X_n\}$ be a submartingale satisfying

$$supE[|X_n|] < \infty$$

Then there exists a random variabe X_∞ to which $\{X_n\}$ converges with probability one,

$$Pr\{\lim_{n\to\infty} X_n = X_\infty\} = 1$$

- If $\{X_n\}$ is a martingale and is uniformly integrable then, in addition to what said before, $\{X_n\}$ converges in the mean, that is,

$$\lim_{n\to\infty} E[|X_n - X_\infty|] = 0$$

and

$$E[X_\infty] = E[X_n]$$

for all n

Chapter 2

Martingale Pricing Theory

Martingale pricing is a pricing approach based on the notions of martingale and risk neutrality. The martingale pricing approach is a cornerstone of modern quantitative finance and can be applied to a variety of derivatives contracts, e.g. options, futures, interest rate derivatives, credit derivatives, etc.

2.1 Introduction to options ,securities ,state prices,single and multi period model

- options:

  In finance, an option is a contract which gives the buyer (the owner or holder of the option) the right, but not the obligation, to buy or sell an underlying asset or instrument at a specified strike price on a specified date, depending on the form of the option. The strike price may be set by reference to the spot price (market price) of the underlying security or commodity on the day an option is taken out, or it may be fixed at a discount or at a premium. The seller has the corresponding obligation to fulfill the transaction – to sell or buy – if the buyer (owner) "exercises" the option. An option that conveys to the owner the right to buy at a specific price is referred to as a call; an option that conveys the right of the owner to sell at a specific price is referred to as a put. Both are commonly traded, but the call option is more frequently discussed.

The seller may grant an option to a buyer as part of another transaction, such as a share issue or as part of an employee incentive scheme, otherwise a buyer would pay a premium to the seller for the option. A call option would normally be exercised only when the strike price is below the market value of the underlying asset, while a put option would normally be exercised only when the strike price is above the market value. When an option is exercised, the cost to the buyer of the asset acquired is the strike price plus the premium, if any. When the option expiration date passes without the option being exercised, then the option expires and the buyer would forfeit the premium to the seller. In any case, the premium is income to the seller, and normally a capital loss to the buyer.

The owner of an option may on-sell the option to a third party in a secondary market, in either an over-the-counter transaction or on an options exchange, depending on the option. The market price of an American-style option normally closely follows that of the underlying stock, being the difference between the market price of the stock and the strike price of the option. The actual market price of the option may vary depending on a number of factors, such as a significant option holder may need to sell the option as the expiry date is approaching and does not have the financial resources to exercise the option, or a buyer in the market is trying to amass a large option holding. The ownership of an option does not generally entitle the holder to any rights associated with the underlying asset, such as voting rights or any income from the underlying asset, such as a dividend.

- security:

  A security is a tradable financial asset. The term commonly refers to any form of financial instrument, but its legal definition varies by jurisdiction. In some jurisdictions the term specifically excludes financial instruments other than equities and fixed income instruments. In some jurisdictions it includes some instruments that are close to equities and fixed income, e.g., equity warrants. In some countries and languages the term "security" is commonly used in day-to-day parlance to mean any form of financial instrument, even though the underlying legal and regulatory regime may not have such a broad definition.

- state price security:

  A state-price security, also called an Arrow-Debreu security (from its origins in the Arrow-Debreu model), a pure security, or a primitive security is a contract that agrees to pay one unit of a numeraire (a currency or a commodity) if a particular state occurs at a particular time in the future and pays zero numeraire in all the other states. The price of this security is the state price of this particular state of the world. The state price vector is the vector of state prices for all states.As such, any derivatives contract whose settlement value is a function of an underlying asset whose value is uncertain at contract date can be decomposed as a linear combination of its Arrow-Debreu securities, and thus as a weighted sum of its state prices.

The Arrow-Debreu model (also referred to as the Arrow-Debreu-McKenzie model or ADM model) is the central model in general equilibrium theory and uses state prices in the process of proving the existence of a unique general equilibrium.

- single and multi period model:

  Single Period Model, one of the discounted cash flow models, is an income valuation approach that aims to find the fair value of a stock/firm using single projected cash flow value and then discounting it with an appropriate discount rate. Taking all future streams of cash flow into one single period and discounting is also referred as "Earnings Capitalisation".This method is a substitute for the traditional discounting of all future cash flows. However since it is a "single period" model, we need a single sum of an amount as the cash flow for all future years or a single sum for 1 year holding period.Formula for calculating value of firm/company using single period model:

  Value of a firm or company= Net Income / Discounting Rate

  Let us take an example:

  To estimate the value of the firm, company or project, stabilized net operating income is divided by an appropriate discount rate. Assuming a stable earning (net of expenses) of USD 300,000 per annum and a discount rate of 12%, the value of the firm can be calculated as follows:

Value = Net Income / Discounting Rate

= $ 300,000 / 0.12

=$ 2,500,000

If a growth number needs to be adjusted to the model, assuming a constant growth of 5%, the value of the firm can now be calculated as follows:

Value = Net Income / Discounting Rate

= $ 300,000 / (0.12 -0.05)

= $ 300,000 / 0.07

=$ 4,285,714

When the discount rate and growth rate are assumed to remain constant from day of valuation till perpetuity, the single period model will yield same results as multi period model.

The single period method of valuation is best suited in case of stable net income flows or cases where it is extremely difficult to forecast future series of cash flows or in cases where the holding period of the investment is 1 year. Selecting the appropriate discount rate may, however, remain a challenging task and would entail estimation error.

For limitations faced with single period error; the improved model, which involves using multiple cash flow forecasting and discounting them, is used with the intent of reducing the estimation error. The said model is also known as Multi-Period Discounted Cash Flow Model.

A discounted cash flow (DCF) analysis is a method of valuing a project, company, or asset using the concepts of the time value of money. All future cash flows are estimated and discounted by using cost of capital to give their present values (PVs). The sum of all future cash flows, both incoming and outgoing, is the net present value (NPV), which is taken as the value of the cash flows in question. Using DCF analysis to compute the NPV takes as input cash flows and a discount rate and gives as output a present value; the opposite process—takes cash flows and a price (present value) as inputs, and provides as output the discount rate—this is used in bond markets to obtain the yield.

The discounted cash flow formula is derived from the future value formula for calculating the time value of money and compounding returns.

$$DCF = \frac{CF_1}{(1+r)^1} + ... + \frac{CF_n}{(1+r)^n}$$

$$FV = DCF \times (1+r)^n$$

where:

- DPV is the discounted present value of the future cash flow (FV), or FV adjusted for the delay in receipt;
- FV is the nominal value of a cash flow amount in a future period;
- r is the interest rate or discount rate, which reflects the cost of tying up capital and may also allow for the risk that the payment may not be received in full;
- n is the time in years before the future cash flow occurs.

Where multiple cash flows in multiple time periods are discounted, it is necessary to sum them as follows:

$$DVP = \sum_{t=0}^{N} \frac{FV_t}{(1+r)^t}$$

for each future cash flow (FV) at any time period (t) in years from the present time, summed over all time periods. The sum can then be used as a net present value figure. If the amount to be paid at time 0 (now) for all the future cash flows is known, then that amount can be substituted for DPV and the equation can be solved for r, that is the internal rate of return. All the above assumes that the interest rate remains constant throughout the whole period. If the cash flow stream is assumed to continue indefinitely, the finite forecast is usually combined with the assumption of constant cash flow growth beyond the discrete projection period. The total value of such cash flow stream is the sum of the finite discounted cash flow forecast and the Terminal value (finance).

For continuous cash flows, the summation in the above formula is replaced by an integration:

$$DVP = \int_0^T FV(t) \exp^{-} \lambda t dt$$

$$= \int_0^T \frac{FV(t)}{(1+r)^t} dt$$

where FV(t) is now the rate of cash flow, and $\lambda = log(1 + r)$.

2.2 Arbitrage pricing theory

Arbitrage pricing theory is an asset pricing model based on the idea that an asset's returns can be predicted using the relationship between that asset and many common risk factors. Created in 1976 by Stephen Ross, this theory predicts a relationship between the returns of a portfolio and the returns of a single asset through a linear combination of many independent macroeconomic variables.

The arbitrage pricing theory (APT) describes the price where a mispriced asset is expected to be. It is often viewed as an alternative to the capital asset pricing model (CAPM), since the APT has more flexible assumption requirements. Whereas the CAPM formula requires the market's expected return, APT uses the risky asset's expected return and the risk premium of a number of macroeconomic factors. Arbitrageurs use the APT model to profit by taking advantage of mispriced securities, which have prices that differ from the theoretical price predicted by the model. By shorting an overpriced security, while concurrently going long in the portfolio the APT calculations were based on, the arbitrageur is in a position to make a theoretically risk-free profit.

- Arbitrage pricing theory equation and example:

  APT states that the expected return on a stock or other security must adhere to the following relationship: Expected return = $r(f) + b(1) \times rp(1) + b(2) \times rp(2) + ... + b(n) \times rp(n)$ Where,

  r(f) = the risk-free interest rate

  b = the sensitivity of the asset to the particular factor

rp = the risk premium associated with the particular factor

The number of factors will range depending on the analysis. There can be a few or dozens; it depends on which factors an analyst chooses for the analysis. In addition, the exact factors do not have to be the same across analyses. As an example calculation, assume a stock is being analyzed. The following four factors have been identified, along with the stocks sensitivity to each factor and the risk premium associated with each factor:

Gross domestic product growth: b = 0.6, rp = 4%

Inflation rate: b = 0.8, rp = 2%

Gold prices: b = -0.7, rp = 5%

Standard and Poor's 500 index return: b = 1.3, rp = 9%

The risk-free rate is 3%

Using the above APT formula, the expected return is calculated as: Expected return $= 3\% + (0.6 \times 4\%) + (0.8 \times 2\%) + (-0.7 \times 5\%) + (1.3 \times 9\%) = 15.2\%$

2.3 The Black Scholes option pricing theory

The Black-Scholes formula (also called Black-Scholes-Merton) was the first widely used model for option pricing. It's used to calculate the theoretical value of European-style options using current stock prices, expected dividends, the option's strike price, expected interest rates, time to expiration and expected volatility.

The Black-Scholes model makes certain assumptions:

- The option is European and can only be exercised at expiration.
- No dividends are paid out during the life of the option.
- Markets are efficient (i.e., market movements cannot be predicted).
- There are no transaction costs in buying the option.
- The risk-free rate and volatility of the underlying are known and constant.
- The returns on the underlying are normally distributed.

$$C = SN(d_1) - N(d_2)K\exp^{-rt}$$

$$d_1 = \frac{ln(\frac{S}{K}) + (r + \frac{s^2}{2})t}{s\sqrt{t}}$$

$$d_2 = d_1 - s\sqrt{t}$$

C=Call premium

S=Current stock price

t=Time until option exercise

K=Option striking price

r=Risk free interest rate

N=Cumulative standard normal distribution

s=Standard derivation

ln=Natural log

The model is essentially divided into two parts: the first part,$SN(d_1)$, multiplies the price by the change in the call premium in relation to a change in the underlying price. This part of the formula shows the expected benefit of purchasing the underlying outright. The second part,$N(d_2)K\exp^{-rt}$, provides the current value of paying the exercise price upon expiration (remember, the Black-Scholes model applies to European options that can be exercised only on expiration day). The value of the option is calculated by taking the difference between the two parts, as shown in the equation.

Chapter 3

Testing the Martringale Hypothesis

This chapter examines testing the Martingale Difference Hypothesis (MDH) and related statistical inference issues. The earlier literature on testing the MDH was based on linear measures of dependence, such as sample autocorrelations, for instance the classical Box-Pierce Portmanteau test and the Variance Ratio test. In order to account for the existing nonlinearity in economic and financial data, two directions have been entertained. First, to modify these classical approaches by taking into account the possible nonlinear dependence. Second, to use more sophisticated statistical tools such as those based on empirical processes theory or the use of generalized spectral analysis. This chapter discusses these developments and applies them to exchange rate data.

3.1 Introduction

Martingale testing has historically received an enormous attention in econometrics. One of the main reasons is the efficient market hypothesis (MDH) and the many ideas related to it. In addition, many economic theories in dynamic contexts in which expectations are assumed to be rational lead to such dependence restrictions on the underlying economic variables; see e.g. Hall (1978), Fama (1991), LeRoy (1989), Lo (1997) and Cochrane (2005). These have prompted a vast research in macro and financial economics which have stimulated a huge interest in developing suitable econometric techniques. This econometric research has grown around the theme of lack of predictability of macro or financial series, but this topic has flourished in different branches, emphasized different methodological aspects, and appeared under different subject names.

When looking at assets prices, the idea of lack of predictability has been commonly referred to as the random walk hypothesis. Unfortunately, the term random walk has been used in different contexts to mean different statistical objects. For instance, in Campbell, Lo and MacKinlay (1997) textbook, they distinguish three types of random walks according to the dependence structure of the increment series. Random walk 1 corresponds to independent increments, random walk 2 to conditional mean independent increments, and random walk 3 to uncorrelated increments. Of these three notions, the two relevant for financial econometrics are the second and the third. The notion of random walk 1 is clearly rejected in financial data for many reasons, the most important is the volatility, that is, the lack of constancy of the variance of current asset returns conditional on lagged asset returns. Within this terminology, this chapter will focus basically on the idea of random walk 2, but we will also discuss some aspects associated to random walk 3. A martingale would correspond to random walk 2, and it plainly means that the best forecast of tomorrows asset price is todays. Then, the asset returns, which are unpredictable, are said to form a martingale difference sequence. Since asset prices are not stationary, from a technical point of view, it is simpler to consider asset returns. Testing prices follow a martingale, it is more common to test that returns follow a martingale difference sequence.

3.2 The Martingale difference hypothesis or MDH

The Martingale Difference Hypothesis (MDH) plays a central role in economic models where expectations are assumed to be rational. The underlying statistical object of interest is the concept of a martingale or, alternatively, the concept of martingale difference sequence (MDS).

The MDH slightly generalizes the notion of MDS by allowing the unconditional mean of Y_t to be nonzero and unknown. The MDH states that the best predictor, in the sense of least mean square error, of the future values of a time series given the past and current information set is just the unconditional expectation. The MDH is called conditional mean independence in the statistical literature, and it means that past and current information are of no use to forecasting future values of a MDS.

Let $I_t = \{Y_t, Y_{t-1}, ...\}$ be the information set at time t and let F_t be the σ field generated by I_t. Then, the following equivalence is fundamental because it formalizes the characteristic property of a MDS, that is, the fact that Y_t is linearly unpredictable given any linear or nonlinear transformation of the past $\omega(I_{t-1})$. That is

$$E[Y_t|I_{t-1}] = \mu \iff E[(Y_t - \mu)\omega(I_{t-1})] = 0 \qquad (3.2.1)$$

$\mu \in R$ for all F_{t-1} measurable weighting function $\omega(.)$ (such that the moment exists). Equation (3.2.1) is fundamental to understand the motivation and main features behind many tests for the MDH. There are two challenging features present in the definition of a mds: first, the information set at time t, I_t, will typically include the infinite past of the series, and second, the number of functions $\omega(.)$ is also infinite. We will classify the extant theoretical literature on testing the MDH, according to what types of functions $\omega(.)$ are employed. We shall illustrate some of the available methods for testing the MDH by applying them to exchange rate returns. The martingale properties of exchange rate returns have been studied previously by many authors leading to mixed conclusions. For instance, Bekaert and Hodrick (1992), Escanciano and Velasco (2006a, 2006b), Fong and Ouliaris (1995), Hong and Lee (2003), Kuan and Lee (2004), LeBaron (1999), Levich and Thomas (1993), Liu and He (1991), McCurdy and Morgan (1988) and Sweeney (1986) find evidence against the MDH for nominal or real exchange rates at different frequencies, whereas Diebold and Nason (1990), Fong, Koh and Ouliaris (1997), Hsieh (1988, 1989, 1993), McCurdy and Morgan (1987) and Meese and Rogo(1983a,b) find little evidence against the MDH. Here, we consider data that consists of four daily and weekly exchange rate returns on the Euro (Euro), Canadian Dollar (Can), the sterling Pound (pound) and the Japanese Yen (yen) against the US dollar. The daily data is taken from January 2, 2004 to August 17, 2007, with a total of 908 observations. As for the weekly data, we consider the returns on Wednesdays from January 2, 2000 to August 17, 2007, with a total of 382 observations. The daily noon buying rates in New York City certified by the Federal Reserve Bank of New York for customs and cable transfers purposes are obtained from http://www.federalreserve. gov/Releases/h10/hist. In figure 3.1 and figure 3.2 we have plotted the evolution of these four daily series for the whole period from January 2, 2000 to August 17, and again similarly to previous analysis, the main two features of these plots are their unpredictibality and their volatility.

Table 3.1 provides summary statistics for the most relevant aspects of the marginal distribution of the data. Similarly to most financial series the main feature from Table 3.1 is the kurtosis(measure of the "peaklness" of the probability distribution of a real-valued random variable) that, in the line of previous studies, is larger for daily than for weekly data. Note that skewness (measure of the asymmetry of the probability distribution of a real-valued random variable about its mean) is moderate and slightly negative for daily data. As it has been observed repeatedly before, the marginal distribution of weekly data is closer to the normal distribution than that of daily data.

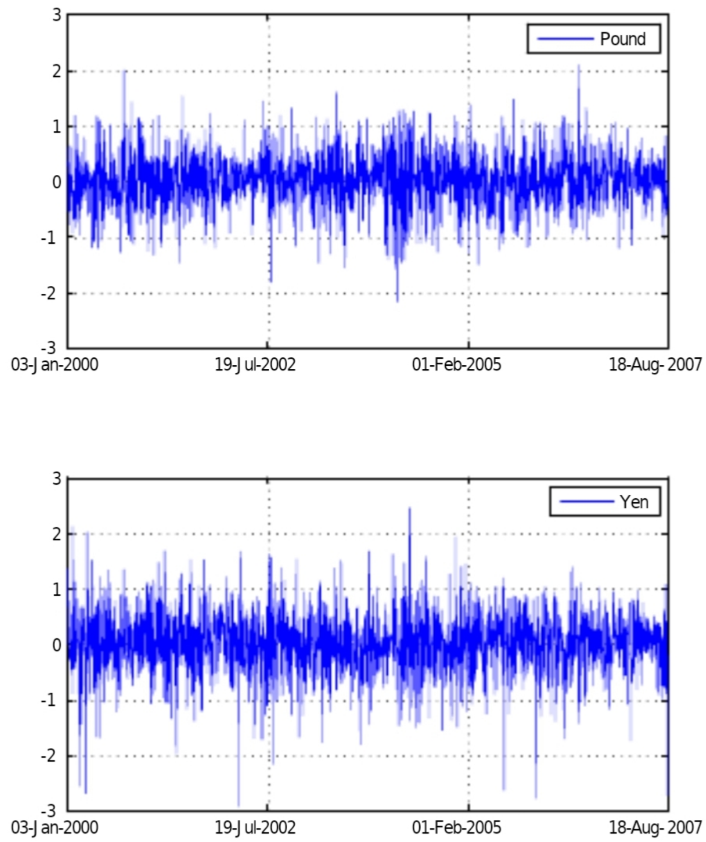

Figure 3.1: **Daily returns of the sterling pound (Pound) and the Japanese Yen (Yen) against the US dollar.**

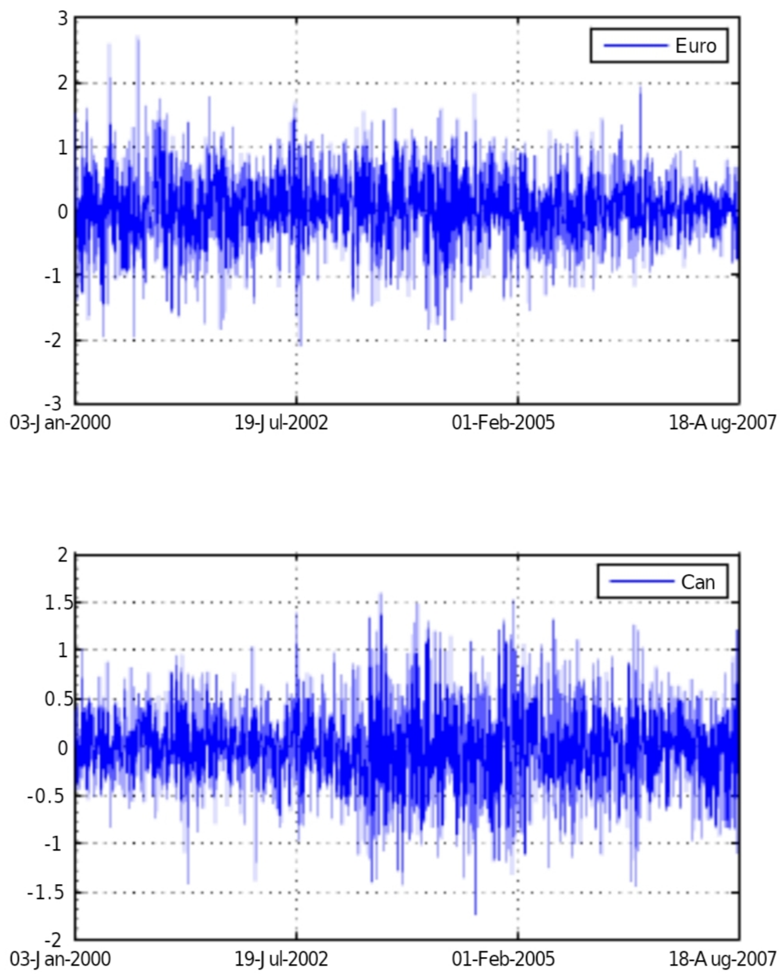

Figure 3.2: **Daily returns of the Euro (Euro), Canadian Dollar (Can) against the US dollar.**

Table 3.1: **Summary statistics of exchange rates returns**

	Daily				Weekly			
	Euro	**Pound**	**Can**	**Yen**	**Euro**	**Pound**	**Can**	**Yen**
n	908	908	908	908	382	382	382	382
Mean	0.0076	0.0113	-0.0213	0.0068	0.0738	0.0552	-0.0832	0.0352
Median	0.0000	0.0221	-0.0080	0.0279	0.0781	0.0763	-0.0864	0.0141
SD	0.5423	0.5332	0.5036	0.5670	1.3539	1.1407	0.9410	1.2525
Skewness	-0.1263	-0.0976	-0.0196	-0.3763	0.0540	0.0545	0.0846	-0.2945
Kurtosis	3.7602	3.4927	3.1345	5.0746	3.0555	2.9649	2.8875	3.0895
Maximum	1.9358	2.0930	1.5129	2.4519	4.4680	3.4830	2.8128	3.1835
Minimum	-2.0355	-2.1707	-1.7491	-2.7859	-3.1636	-3.2307	-2.7067	-4.3058

3.3 Tests based on linear measures of dependence

Recall the MDS denition in equation (3.1) that should hold for any function $\omega(.)$. The simplest approach is to consider linear functions $\omega(.)$, such as $\omega(I_{t-1}) = Y_{t-j}$, for some $j \geq 1$. Hence, a necessary (but not sufficient, in general) condition for the the MDH to hold is that the time series is uncorrelated, i.e.

$$\gamma_j = cov(Y_t, Y_{t-j}) = E[(Y_t - \mu)Y_{t-j}] = 0 \qquad (3.3.1)$$

for all $j \geq 1$ where γ_j denotes the autocovariance of order j. In principle, one should test that all autocovariances or autocorrelations are zero. However, the most employed tests just consider that a finite number of autocorrelations are zero.

Notice that the early literature, which includes some distinguished references such as Yule (1926), Bartlett (1955), Grenander and Rosenblatt (1957) or Durbin and Watson (1950), essentially assumed Gaussianity and, hence, identified three concepts: lack of serial correlation, MDS and independence. In the time series literature the term white noise is commonly used to denote an uncorrelated series that can present some form of dependence. Obviously, a white noise series is not necessarily independent nor MDS since dependence can be reflected in other aspects of the joint distribution such as higher order moments. The distinction between these three concepts has been stressed recently in econometrics. In fact, during the past years a variety of models designed to reflect nonlinear dependence has been studied in the econometrics literature. For instance, in empirical finance, ARCH (Auto Regressive Conditional Hetrosketastic Mode)l and bilinear models have been widely studied, see Bera and Higgins (1993, 1997) and Weiss (1986) for a comparison. These models are suitable to reflect the nonlinear dependence structure found in many financial series.

Tests for white noise have been proposed both in the time domain and in the frequency domain. The time domain has mainly, but not exclusively, focused on a finite number of lags, while the frequency domain has been more suitable to address the infinite dimensional case.

3.3.1 Tests based on a finite dimensional conditioning set

In the time domain the most popular test has been the Box-Pierce (Box and Pierce, 1970) Portmanteau Q_p test. The Q_p test is designed for testing that the first p autocorrelations of a series (possibly residuals) are zero. The number p can be considered to be fixed or to grow with the sample size n. In this section we will assume that p is fixed.

Suppose that we observe raw data $\{Y_t\}_{t=1}^n$. Then, γ_j can be consistently estimated by the sample autocovariance

$$\hat{\gamma_j} = \frac{\sum_{t=1+j}^n (Y_t - \bar{Y})(Y_{t-j} - \bar{Y})}{(n-j)}$$

where $\bar{Y}$ is the sample mean, and we also introduce $\hat{\rho} = \hat{\gamma_j}/\hat{\gamma_0}$ to denote the j-th order autocorrelation. the Q_p statistic is just

$$Q_p = n \sum_{j=1}^p \hat{\rho}_j^2$$

but it is commonly implemented via the Ljung and Box (1978) modification

$$LB_p = \frac{n(n+2)\sum_{j=1}^p \hat{\rho}_j^2}{(n-j)}$$

Note that Q_p (orLB_p) only takes into account the linear dependence up to the lag p: When p is considered fixed, the Q_p test statistic applied to independent data follows assymptotically χ_p^2 distribution since the asymptotic covariance matrix of the first p autocorrelations of an independent series is the identity matrix. Hence, it is useful to write $Q_p = (\sqrt{n}\underline{\hat{\rho}})' I^{-1} (\sqrt{n}\underline{\hat{\rho}})$ where $\underline{\hat{\rho}} = (\underline{\hat{\rho}}_1, ..., \underline{\hat{\rho}}_p)'$. Note, however, that when the series present some kind of nonlinear dependence this asymptotic null covariance matrix is no longer the identity. In fact, denoting $\underline{\rho} = (\underline{\rho}_1, ..., \underline{\rho}_p)'$ for a general time series the asymptotic distribution of $\sqrt{n}(\underline{\hat{\rho}} - \underline{\rho})$ is N(0,T) where the (i,j)-th element of T is given by

$$\gamma_0^{-2}(c_{ij} - \rho_i \times c_{0j} - \rho_j \times c_{0i} + \rho_i \times \rho_j \times c_{00})$$

where, for $i, j = 0, 1, ..., p$, $c_{ij} = \sum_{d=-\infty}^{\infty}\{E[(Y_t-\mu)(Y_{t-i}-\mu)(Y_{t+d}-\mu)(Y_{t+d-j}-\mu)] - E[(Y_t-\mu)(Y_{t-i}-\mu)]E[(Y_{t+d}-\mu)(Y_{t+d-j}-\mu)]$

Under alternative assumptions the matrix T can be simplified and this will lead to several modified versions of the Box-Pierce statistic. When this matrix is still diagonal, as it happens under MDS and additional moment restrictions, which, for instance, are satifised by Gaussian GARCH models and many stochastic volatility models, the natural approach is to robustify the Q_p by standardizing it by a consistent estimation of its asymptotic variance, i.e.,

$$Q_p^* = n \sum_{j=1}^{p} \frac{\hat{\rho}_j^2}{\tau_j}$$

where

$$\tau_j = \hat{\gamma_0}^{-2} \sum_{t=1+j}^{n} (Y_t - \bar{Y})^2 (Y_{t-j} - \bar{Y})^2$$

We have followed Lobato, Nankervis and Savin (2001)notation and denoted Q_p by Q_p^*. This statistic has appeared in different versions, see for instance, Diebold (1986), Lo and MacKinlay (1989), Robinson (1991), Cumby and Huizinga (1992), Bollerslev and Wooldridge (1992) and Bera and Higgins (1993). The Q_p^* statistic (or its Ljung-Box analog) should be routinely computed for financial data instead of the standard Q_p (or the LB_p). However, this is not typically the case. For the general case, the asymptotic covariance matrix of the first p autocorrelations is not a diagonal matrix. Hence, for this general case both the Q_p and the Q_p^* tests are invalid. However, under MDS the matrix T can be greatly simplified so that its ij-th element takes the form $E[(Y_t-\mu)^2(Y_{t-i}-\mu)(Y_{t-j}-\mu)]$ that can be easily estimated using its sample analog. This is the approach followed by Guo and Phillips (2001). For the general case, that includes MDS and non MDS processes, the asymptotic covariance matrix of the first p autocorrelations is a complicated nondiagonal matrix. Hence, for this general case, the literature has proposed the following two modifications of the Q_p test. The first one is to modify the Q_p statistic by introducing a consistent estimator of the asymptotic null covariance matrix of the sample autocorrelations $\hat{\tau}$ so that the modified Q_p statistic retains the asymptotically χ_p^2 null distribution. Lobato, Nankervis and Savin (2002) name this statistic

$$\tilde{Q}_p = (\sqrt{n}\hat{\rho})'\hat{\tau}^{-1}(\sqrt{n}\hat{\rho})$$

The main drawback of this approach is that in order to construct $\hat{\tau}$ a bandwidth number has to be introduced. This approach works for general dependence structures that allow for the asymptotic covariance matrix of the first p autocorrelations to take any form. The second modification has been studied by Horowitz, Lobato, Nankervis and Savin (2006) who employ a bootstrap procedure to estimate consistently the asymptotic null distribution of the Q_p test for the general case. They compare two bootstrap approaches, a single and a double blocks-of-blocks bootstrap, and the final recommendation is to employ a double blocks-of-blocks bootstrap after prewhitening the time series. This solution presents a similar problem, though, namely the researcher has to choose arbitrarily a block length number. The previous papers considered raw data, but Francq, Roy and Zakoan (2005) have addressed the use of the Q_p statistic with residuals. They propose to estimate the asymptotic null distribution of the Q_p test statistic for the general weak dependent case. However, their approach still requires the selection of p, and of several additional arbitrary numbers necessary to estimate consistently the needed asymptotic critical values.

These previous references represent an effort to address the problem of testing for mds using the standard linear measures (autocorrelations) but allowing for nonlinear dependence. Lobato (2001) represents an alternative approach with a similar idea. The target is to avoid the problem of introducing a user-chosen number and the idea is to construct an asymptotically distribution free statistic. Although this approach delivers tests that handle nonlinear dependence and control properly the type I error in finite samples, its main theoretical drawback is its inefficiency in terms of local power.

A related statistic, which has been commonly employed in the empirical finance literature is the variance ratio that takes the form

$$VR_p = 1 + 2\sum_{j=1}^{p-1}(1-(j/p))\hat{\rho}_j$$

under independence, $\sqrt{np}(VR_p - 1)$ is asymptotically distributed as N(0,2(p-1)). Although this test can also be robustified and it can be powerful in some occasions, it presents the serious theoretical limitation of being inconsistent. For instance, Gonzalez and Lobato (2003) considered an MA(2): $y_t = e_t - 0.4597e_{t-1} + 0.10124e_{t-2}$.For this process $VR_3 = 0$ in spite that the first two autocorrelations are nonzero. The problem with variance ratio statistics resides in the possible existence of compensations between autocorrelations with different signs, and this may affect power severely. Related to VR tests, Nankervis and Savin (2007) have proposed a robustified version of the

Andrews and Plobergers (1996) test that appear to have very good finite sample power with the common empirical finance models. Also related, Delgado and Velasco (2007) have recently considered a large class of directional tests based on linear combinations of autocorrelations. Their tests are shown to be optimal in certain known local alternative directions and are asymptotically equivalent to Lagrange Multiplier tests. Finally, we mention Kuan and Lee (2004) who propose a correlation-based test for the MDH that instead of using lagged values of Y_t as the function $\omega(.)$, they employ some other arbitrary $\omega(.)$. This test shares with all the tests analyzed in this section the problem of inconsistency derived from not using a whole family of functions $\omega(.)$.

3.3.2 tests based on a infinite dimensional conditioning set

The approach presented in the previous subsection laid naturally in the time domain since a finite number of autocorrelations were tested. However, when the infinite past is considered, the natural framework for performing inference is the frequency domain. The advantage of the frequency domain is the existence of one object, namely, the spectral density, that contains the information about all the autocovariances. Hence, in the frequency domain, the role previously taken by autocorrelations is now carried out by the spectral density function. The spectral density $f(\lambda)$ is defined implicitly by

$$\gamma_k = 1/2\pi \int_{-\pi}^{\pi} f(\lambda) exp(ik\lambda) d\lambda$$

for k=0,1,2...

Define also the periodogramas $I(\lambda) = |\omega(\lambda)|^2$ where $\omega(\lambda) = n^{-1/2} \sum_{t=1}^{n} x_t exp(it\lambda)$ Although the periodogram is an inconsistent estimator of the spectral density, it can be employed as a building block to construct a consistent estimator. The integral of the spectral density is called the spectral distribution, which, under the MDH, is linear in λ .

For this infinite lag case, the MDH implies as null hypothesis of interest that $\gamma_k = 0$ for all $k \neq 0$, and equivalently, in terms of the spectral density, the null hypothesis states that $f(\lambda) = \gamma_0/2\pi$ for all $\lambda \in [-\pi;\pi]$.

The advantage of the frequency domain is that the problem of selecting p, which was present in the previous subsection, does not appear because the null hypothesis is stated in terms of all autocorrelations, as summarized by the spectral density or distribution. The classical approach in the frequency domain involves the standardized cumulative periodogram, that is,

$$Z_n = \sqrt{T}(\frac{\sum_{j=1}^{[\lambda T/\pi]} I(\lambda_j)}{\sum_{j=1}^{T} I(\lambda_j)} - \frac{\lambda}{\pi})$$

where $j = 1, 2...n/2$ are called the Fourier frequencies. Based on $Z_n(\lambda)$, the two classical tests statistics are the Kolmogorov-Smirnov $\max_{j=1,...,T} |Z_n(\lambda_j)|$ and the Cramer von Mises $T^{-1} \sum_{j=1}^{T} Z_n(\lambda_j)^2$.

These tests statistics have been commonly employed (see Bartlett (1955) and Grenander and Rosenblatt(1957)) because when the series y_t is not only white noise but also independent (or MDS with additional moment restrictions), it can be shown that the process $Z_n(\lambda)$ converges weakly in $D[0;\pi]$ (the space of right continuous functions in $D[0;\pi]$) to the Brownian bridge process, see

Dahlhaus (1985). Hence, asymptotic critical values are readily available for the independent case. In fact, Durlauf (1991) has shown that the independence assumption can be relaxed to conditional homoskedastic mds. For the mds case with conditional heteroskedasticity (and some moment conditions), Deo (2000) slightly modified this statistic so that the standardized cumulative periodogram retained the convergence to the Brownian bridge. Deo's(2000) test can be interpreted as a continuous version of the robustified Box-Pierce statistic, Q_p^* Notice that in Deo's setup there is no need of introducing any user-chosen number since under the stated assumptions the autocorrelations are asymptotically independent. As Deo comments, his assumption is the main responsible for the diagonality of the asymptotic null covariance matrix of the sample autocorrelations. However, for many common models, such as GARCH models with asymmetric innovations, EGARCH models and bilinear models, Deo's condition does not hold and the autocorrelations are not asymptotically independent under the null hypothesis. Hence, for the general case, Deo's test is not asymptotically valid. Deo's Cramer-von Mises test statistic can also be written in the time domain as

$$DEO_n = \sum_{j=1}^{n-1} n \frac{\hat{\rho}_j^2}{\tau_j} (\frac{1}{j\pi})^2$$

More general weighting schemes for the sample autocovariances $\hat{\rho}_j$ than the ones considered here are possible. Under the null hypothesis of the MDS and some additional assumptions

$$DEO_n \to_d \int_0^1 B^2(t)dt$$

as $n \to \infty$

where B(t) is the standard Brownian bridge on [0,1]. The 10%, 5% and 1% asymptotic critical values can be obtained from Shorack and Wellner and are 0.347, 0.461 and 0.743, respectively.

Chen and Romano (1999, p.628) estimated the asymptotic distribution by means of either the block bootstrap or the subsampling technique. Unfortunately, these bootstrap procedures require the selection of some arbitrary number and in a general framework no theory is available about their optimal selection. Alternative bootstrap procedures which do not require the selection of a user chosen number such as resampling the periodogram as in Franke and Hardle (1992) or in Dahlhaus and Janas (1996) will not estimate consistently the asymptotic distribution because of the fourth order cumulant terms.

Lobato and Velasco (2004) considered the use of the statistic

$$M_n = \frac{T^{-1}\sum_{j=1}^{T} I(\lambda_j)^2}{[T^{-1}\sum_{j=1}^{T} I(\lambda_j)]^2} - 1$$

under general weak dependence conditions. This statistic was previously considered by Milhoj (1981) who employed M_n as a general goodness of fit test statistic for time series. Milhoj informally justified the use of this statistic for testing the adequacy of linear time series models, but since he identified white noise with i.i.d. , his analysis does not automatically apply in general contexts.

Beran (1992) and Deo and Chen (2000) have also employed the M_n test statistic as a goodness-of-fit tests for Gaussian processes. Statistical inference is especially simplified with M_n since its asymptotic null distribution is normal even after parametric estimation. We note that the continuous version of M_n can be expressed in the time domain as an statistic proportional to $\sum_{j=0}^{n-1} \hat{\rho}_j^2$ which shows the difficulty of deriving the asymptotic properties in the time domain since the $\hat{\rho}_j$ may not be asymptotically independent.

In the time domain, Hong (1996) has considered p as growing with n and hence, has been able to derive a consistent test in the time domain for the case of regression residuals. In this framework p can be interpreted as a bandwidth number that needs to grow with n so that his test can handle the fact that the null hypothesis implies an infinite number of autocovariances. Hong restricted to the independent case while Hong and Lee (2003) have extended Hongs procedure to allow for conditional heteroskedasticity. However, notice that their framework still restricts the sample autocorrelations to be asymptotically independent. An alternative solution recently explored by Escanciano and Lobato (2007) consists on modifying the Box-Pierce statistic using an adaptive Neyman test that would take the form $N_n = Q_{\tilde{p}}^*$ where $\tilde{p} = min\{m; 1 \leq m \leq p_n; L_m \geq L_h, h = 1, ...p_n\}$ where $L_p = Q_p^* - \pi(p, n, q)$ and p_n is an upper bound that grows slowly to infinite with n, and

$$\pi(p, n, c) = plogn, if \max_{1 \leq j \leq p_n} |\frac{\hat{\rho}_j^2}{\tau_j}| \leq \sqrt{qlogn}$$

$$\pi(p, n, c) = 2p, if \max_{1 \leq j \leq p_n} |\frac{\hat{\rho}_j^2}{\tau_j}| > \sqrt{qlogn}$$

where q is some fixed positive number. Our choice of q is 2.4 and it is motivated from an extensive simulation study in Inglot and Ledwina (2006) and from simulations in Escanciano and Lobato (2007). Small values of q result in the Akaikes criterion choice (The Akaike information criterion (AIC) is an estimator of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, AIC provides a means for model selection), while large q's lead to the choice of the Schwarzs criterion(SIC is a criterion for model selection among a finite set of models; the model with the lowest SIC is preferred. It is based, in part, on the likelihood function). Moderate values, such as 2.4; provide a switching effect in which one combines the advantages of the two selection rules, that is, when the alternative is of high frequency (i.e. when only the signicant autocorrelations are at large lags j) Akaike is used whereas if the alternative is of low-frequency (i.e. if the first autocorrelations are different from zero) Schwarz is chosen. The previous adaptive test is an improvement with respect to the traditional Box-Pierce and Hongs approaches because the N_n test is more powerful and less sensitive to the selection of the bandwidth number p_n than these approaches, and more importantly, because it avoids the estimation of the complicated variance-covariance matrix T since its asymptotic distribution is χ^2_1 for general MDS processes.

Summarizing, testing the MDH using linear measures of dependence presents two challenging features. The first aspect is that the null hypothesis implies that an infinite number of autocorrelations are zero. This feature has been addressed successfully in the frequency domain under severe restrictions on the dependence structure of the process. The second feature is that the null hypothesis allows the time series to present some form of dependence beyond the second moments. This dependence entails that the asymptotic null covariance matrix of the sample autocorrelations is not diagonal, so that it has n^2 non-zero terms (contrary to Durlauf (1991) and Deo (2000) who consider a diagonal matrix, and hence, it has only n non-zero elements). This aspect has been handled by introducing some arbitrary user-chosen numbers whose selection complicates statistical inference. However, all these tests are suitable for testing for lack of serial correlation but not necessarily for the MDH, and in fact, they are not consistent against non-martingale difference sequences with zero autocorrelations. This happens when only nonlinear dependence is present, as is commonly the case with financial data, e.g. exchange rates dynamics. These tests are inconsistent because they only employ information contained in the second moments of the process.

To circumvent this problem we could take into account higher order moments, as in Hinich and Patterson (1992). They proposed to use the bispectrum, i.e., the Fourier transform of the third order cumulants of the process, but again, this test is not consistent against non-martingale difference sequences with zero third order cumulants.

In Table 3.2 we report the robust version (to conditional heteroskedasticity) of the first five autocorrelations, the Ljung and Boxs (1978) test, that is a corrected Q_p^* statistic, which we call LB_p^*, Deo's (2000) modification of Durlaufs test statistic and the Escanciano and Lobatos (2007) test based on N_n to check whether or not our exchange rates changes are uncorrelated. This Table is in agreement with previous findings that have shown that exchange rates have no linear dependence, see for instance, Table 2 in Hsieh (1989), Bera and Higgins (1997), Hong and Lee (2003) and references therein.

Table 3.2: **Linear predictability of exchange rates returns**

	Daily				Weekly			
	Euro	**Pound**	**Can**	**Yen**	**Euro**	**Pound**	**Can**	**Yen**
$\hat{\rho}_1$	-0.047	0.001	-0.016	-0.020	0.018	0.046	-0.023	0.054
$\hat{\rho}_2$	0.003	0.007	-0.028	-0.015	-0.002	-0.008	0.031	-0.024
$\hat{\rho}_3$	-0.046	-0.055	-0.001	-0.016	0.049	-0.031	0.011	0.010
$\hat{\rho}_4$	-0.002	0.028	-0.060	0.013	0.024	-0.043	0.015	-0.041
$\hat{\rho}_5$	-0.002	0.003	-0.063	0.039	0.036	-0.024	0.052	-0.095
LB_5^*	4.071	3.586	8.045	2.452	1.795	2.191	1.781	4.900
LB_{15}^*	15.516	13.256	15.181	6.670	9.139	7.451	10.266	18.861
LB_{25}^*	28.552	26.568	19.756	13.155	18.746	32.584	21.786	24.519
LB_{50}^*	61.922	64.803**	49.887	37.428	42.559	59.107	41.140	58.756
N_n	1.889	0.021	0.253	0.380	0.151	0.827	0.208	1.105

Note:* andsignicantly different from zero at the 5% and 10% level, respectively.**

3.4 Tests based on nonlinear measures of dependence

Arguably, testing for the MDH is a challenging problem, since in order to verify it, we must check that a very large class of transformations of the past does not help to predict the current value of the series, see (3.1). An important step through the development of consistent tests was made when the econometricians realized that is not necessary to take a very large class of functions in (3.1) but just a convenient parametric class of functions, satisfying certain properties.This called the integrated approach.

Note, however, that there exists an alternative methodology that is based on the direct estimation of the conditional expectation $E[Y_t|\tilde{Y}_{t,P}]$ where $\tilde{Y}_{t,P} = (Y_{t-1}, ..., Y_{t-P})'$ for some P finite. This approach can be called the smoothing approach or a local approach.Tests within the local approach have been proposed by Wooldridge (1992), Yatchew (1992), Horowitz and Hrdle (1994), Zheng (1996), Fan and Huang (2001), Horowitz and Spokoiny (2001) and Guerre and Lavergne (2005), to mention just a few for a comprehensive review of the local approach when P = 1. Among these tests based on local methods, the test recently proposed by Guay and Guerre (2006) seems to be especially convenient for testing the MDH for two reasons. First, it has been justified for time series under conditional heteroskedasticity of unknown form. Second, it is an adaptive data-driven test (DDT is a software testing methodology that is used in the testing of computer software to describe testing done using a table of conditions directly as test inputs and verifiable outputs as well as the process where test environment settings and control are not hard-coded). Their test combines a chi-square statistic based on nonparametric Fourier series estimators for $E[Y_t|\tilde{Y}_{t,P}]$ coupled with a data-driven choice for the number of components in the estimator. To construct their test a based rank of the unknown conditional variance is needed. Notice that a practically relevant problem of the local approach arises when P is large or even moderate. The problem is motivated by the sparseness of the data in high-dimensional spaces, which leads to most test statistics to suffer a considerable bias, even for large sample sizes. In the next subsection, we will consider an approach that helps to alleviate this problem.

This section focuses on integrated tests. We divide the extensive literature within this integrated approach according to whether the tests consider functions of a finite number of lags or not, that is, whether $\omega(I_{t-1}) = \omega(\tilde{Y}_{t,P})$ for some $P \geq 1$. We stress at the outset that the main advantage of the tests considered in this section is that they are consistent for testing the MDH (at least when the information set has a finite number of variables), contrary to the tests considered in Section 3. The main disadvantage is that their asymptotic null distributions are, in general, not standard, what means that no critical values are ready available. In this situation, the typical solution is to employ the bootstrap to estimate this distribution.

3.4.1 Tests based on a finite dimensional conditioning set

The problem of testing over all possible weighting functions can be reduced to testing the orthogonality condition over a parametric family of functions. Although still the parametric class has to include an infinite number of elements, the complexity of the class to be considered is substantially simplified and makes it possible to test for the MDH.

The methods that we review in this subsection use $\omega(I_{t-1}) = \omega_0(\tilde{Y}_{t,P}, x)$ in (3.1), where $\tilde{Y}_{t,P} = (Y_{t-1}, ..., Y_{t-P})'$ and ω_0 is a known function indexed by a parameter x. That is, these methods check for any form of predictability from the lagged P values of the series. The test statistics are based on a distance from the sample analogue of $E[(Y_t - \mu)\omega_0(\tilde{Y}_{t,P}, x)]$ to zero.

The exponential function $\omega_0(\tilde{Y}_{t,P}, x) = exp(ix'\tilde{Y}_{t,P})$,$x \in R$ was first considered in Bierens (1982, 1984, 1990). One version of the Cramer-von Mises (CvM) test of Bierens (1984) leads to the test statistic

$$CvM_{n,exp,P} =$$

$$\frac{\sum_{t=1}^{n}\sum_{s=1}^{n}(Y_t-\bar{Y})(Y_s-\bar{Y})exp(-1/2|\tilde{Y}_{t,P}-\tilde{Y}_{s,P}|^2)}{n\hat{\sigma}^2}$$

where $\hat{\sigma}^2 = (1/n)\sum_{t=1}^{n}(Y_t-\bar{Y})^2$

Indicator functions $\omega_0(\tilde{Y}_{t,P}, x) = 1(\tilde{Y}_{t,P} \leq x)$, $x \in R$ were used in Stute (1997) and Koul and Stute (1999) for model checks of regressions and autoregressions, respectively, and in Dominguez and Lobato (2003) for the MDH problem. Dominguez and Lobato (2003), extending to the multivariate case the results of Koul and Stute (1999), considered the CvM and Kolmogorov-Smirnov (KS) statistics, respectively,

$$CvM_{n,P} = \frac{\sum_{j=1}^{n}[\sum_{t=1}^{n}(Y_t-\bar{Y})1(\tilde{Y}_{t,P} \leq \tilde{Y}_{j,P})]^2}{\hat{\sigma}^2 n^2}$$

$$KS_{n,P} = \max_{1\leq i\leq n}\frac{|\sum_{t=1}^{n}(Y_t-\bar{Y})1(\tilde{Y}_{t,P} \leq \tilde{Y}_{j,P})|}{|\hat{\sigma}\sqrt{n}|}$$

An important problem of the local approach arises in the case where P is large or even moderate. The sparseness of the data in high-dimensional spaces implies severe biases to most test statistics. This is an important practical limitation for most tests considered in the literature because these biases still persist in fairly large samples. Motivated by this problem, Escanciano (2007a) proposed the use of $\omega_0(\tilde{Y}_{t,P}, x) = 1(\beta'\tilde{Y}_{t,P} \leq u)$ where $x = (\beta, u) \in S^d \times R$,with $S^d = \{\beta \in R^d : |\beta| = 1\}$ and defined CvM tests based on this choice. We denote by $PCVM_{n,P}$ the resulting CvM test in Escanciano (2007a). Also recently, Lavergne and Patilea (2007) has proposed dimension-reduction bootstrap consistent test for regression models based on nonparametric kernel estimators of one-dimensional projections. Their proposal falls in the category of local-based methods, though.

The asymptotic null distribution of integrated tests based on $\omega_0(\tilde{Y}_{t,P}, x)$ depends on the data generating process (DGP) in a complicated way. Therefore, critical values for the tests statistics can not be tabulated for general cases. One possibility, only explored in the literature for the case P = 1 by Koul and Stute (1997), consists of applying the so-called Khmaladzes transformation (Khmaladze, 1981) to get asymptotically distribution free tests. Extensions to $P > 1$ are not available yet. Alternatively, we can approximate the asymptotic null distributions by bootstrap methods. The most relevant bootstrap procedure for testing the MDH has been the wild bootstrap (WB) introduced in Wu (1986) and Liu (1988). For instance, this approach has been employed in Dominguez and Lobato (2003) and Escanciano and Velasco (2006a, 2006b) to approximate the asymptotic distribution of integrated MDH tests. The asymptotic distribution is approximated by replacing $(Y_t - \bar{Y})$ by $(Y_t - \bar{Y})(V_t - \bar{V})$ where $\{V_t\}_{t=1}^n$ is a sequence of independent random variables (rv) with zero mean, unit variance, bounded support and also independent of the sequence $\{Y\}_{t=1}^n$. Here, $\bar{V}$ is the sample mean of $\{V_t\}_{t=1}^n$. The bootstrap samples are obtained resampling from the distribution of V_t. A popular choice for$\{V_t\}$ is a sequence of i.i.d. Bernoulli variates with

$$P(V_t = 0,5(1 - \sqrt{5})) = (1 + \sqrt{5})/2\sqrt{5}$$

and

$$P(V_t = 0,5(1 + \sqrt{5})) = 1 - (1 + \sqrt{5})/2\sqrt{5}$$

We have applied several tests within the integrated methodology to our exchange rates data. In Table 3.3 we report the wild bootstrap empirical values.

Table 3.3: **Testing the MDH of exchange rates returns**

	Daily				**Weekly**			
	Euro	**Pound**	**Can**	**Yen**	**Euro**	**Pound**	**Can**	**Yen**
$CVM_{n,exp,1}$	0.028	0.322	0.744	0.842	0.453	0.086	0.876	0.488
$CVM_{n,exp,3}$	0.164	0.320	0.898	0.666	0.743	0.250	0.076	0.258
$CVM_{n,1}$	0.020	0.354	0.628	0.822	0.610	0.146	0.863	0.388
$CVM_{n,3}$	0.192	0.424	0.798	0.588	0.916	0.893	0.720	0.500
$KS_{n,1}$	0.016	0.220	0.502	0.740	0.726	0.176	0.836	0.542
$KS_{n,3}$	0.036	0.280	0.734	0.526	0.986	0.810	0.224	0.654
$PCVM_{n,1}$	0.020	0.354	0.626	0.822	0.610	0.146	0.863	0.388
$PCVM_{n,3}$	0.248	0.438	0.790	0.664	0.746	0.443	0.566	0.414

In our application we have considered the values P = 1 and P = 3 for the number of lags used in $CVM_{n,exp,P}, CVM_{n,P}, KS_{n,P}$ and $PCVM_{n,P}$.

Our results favor the MDH with all exchange rates at both frequencies, weekly and daily, with the exception of the daily Euro for P = 1. Surprisingly enough, we obtain contradictory results for this exchange rate when P = 3: These contradictory results have been previously documented in e.g. Escanciano and Velasco (2006a) and rather than to a true lack of evidence against the MDH, they may be due to a lack of power of the tests. Although the consideration of an omnibus test, as those discussed in this section, is naturally the first idea when there is no a priori information about directions in the alternative hypothesis, it is worth noting that there is an important limitation of omnibus tests: despite their capability to detect deviations from the null in any direction, it is well-known that they only have reasonable nontrivial local power against very few orthogonal directions, see Janssen (2000) and Escanciano (2008) for theoretical explanations and bounds for the number of orthogonal directions.

Table 3.4: **Testing the MDH of exchange rates returns. Bootstrap P values.Data driven tests**

	Daily				**Weekly**			
	Euro	**Pound**	**Can**	**Yen**	**Euro**	**Pound**	**Can**	**Yen**
$T_{n,\tilde{p}}$	0.049	0.847	0.514	0.876	0.622	0.133	0.747	0.299

A possible solution to overcome the lackof power of omnibus tests is provided by the so-called Neymans smooth tests. They were first proposed by Neyman (1937) in the context of goodness-of-t of distributions, and since then, there has been a lot of research documenting their theoretical and empirical properties. In the context of MDH testing, a recent data-driven smooth test has been proposed by Escanciano and Mayoral (2007). Their test is based on the principal components of the marked empirical processes resulting from the choice $w_0(\tilde{Y}_{t,1}, x) = 1(Y_{t-1} \leq x)$ with $x \in R$: This test is an extension to nonlinear dependence of order one, i.e. for P = 1; of the test based on N_n.As shown by these authors, this test possesses excellent local power properties and compares favorably to omnibus tests and other competing tests. The test statistic is

$$T_{n,\tilde{p}} = \sum_{j=1}^{\tilde{p}} \hat{\epsilon}_{j,n}^2$$

with $T_{n,p}$ replacing Q_p^* there, and where $\hat{\epsilon}_{j,n}$ are the sample principal components of a certain CvM test. The asymptotic null distribution of $T_{n,\tilde{p}}$ is a X_1^2. We have applied the adaptive data-driven test based on $T_{n,\tilde{p}}$ to our exchange rates data. The results are reported in Table 3.4 and support our previous conclusions. Only the MDH for the daily Euro exchange rate is rejected at 1% with $T_{n,\tilde{p}}$

3.4.2 Tests based on an infinite dimensional information set

The afore mentioned references test the MDH conditioning on a finite-dimensional information set, and therefore, they may miss some dependence structure in the conditional mean at omitted lags. In principle, the maximum power could be achieved by using the correct lag order P of the alternative. However, prior information on the conditional mean structure is usually not available.

There have been some proposals considering infinite-dimensional information sets. First, de Jong (1996) generalized Bierens test to time series, and although his test had the appealing property of considering an increasing number of lags as the sample size increases, it required numerical integration with dimension equal to the sample size, which makes this test unfeasible in applications where the sample size is usually large, e.g. financial applications. Second, Dominguez and Lobato (2003) suggest constructing a test statistic as a weighted average of all the tests statistics established for a fixed number of lags. However, Dominguez and Lobato(2003) did not further analyze the test neither the selection of the measure to weight the different statistics.

Using a different methodology based on the generalized spectral density approach of Hong (1999), Hong and Lee (2003) proposed a MDH bootstrap test. Tests based on the generalized spectral density involve three choices: a kernel, a bandwidth parameter and an integrating measure, and, in general, statistical inferences are sensitive to these choices. This fact motivated Escanciano and Velasco (2006a, 2006b) to propose MDH by means of a generalized spectral distribution function.

The generalized spectral approach is based on the fact that the MDH implies that $H_0 : \gamma_{j,w}(x) = 0$,for all x, for all $j \geq 1$. where $\gamma_{j,w}(x) = E[(Y_t - \mu)w_0(Y_{t-j}, x)]$ and where $w_0(Y_{t-j}, x)$ is any of the parametric functions of the previous section. The generalized spectral approach of Hong is based on the choice $w_0(Y_{t-j}, x) = exp(ixY_{t-j})$. Escanciano and Velasco (2006a) considered the latter choice, and Escanciano and Velasco (2006b) used $w_0(Y_{t-j}, x) = 1(Y_{t-j}, x)$, and called the measures $\gamma_{j,ind}(x) = E[(Y_t - \mu)1(Y_{t-j} \leq x)]$ the Integrated Pairwise Autoregression Functions (IPAF). The name follows from the fact that

$$\gamma_{j,ind}(x) = E[(Y_t - \mu)1(Y_{t-j} \leq x)]$$
$$= \int_{-\infty}^{x} E[Y - \mu | Y_{t-j} = z]F(dz)$$

where F is the cdf of Y_t. The measures $\gamma_{j,w}(x)$ can be viewed as a generalization of the usual autocovariances to measure the conditional mean dependence in a nonlinear time series framework. They can be easily estimated from a sample. For instance, the IPAF's can be estimated by

$$\hat{\gamma}_{j,ind}(x) = \frac{\sum_{t=1+j}^{n}(Y_t - \bar{Y})1(Y_{t-j} \leq x)}{n - j}$$

Moreover, as proposed by Escanciano and Velasco (2006b), nonlinear correlograms can be used to formally asses the nonlinear dependence structure in the conditional mean of the series. These authors dene the KS test statistic as

$$KS(j) = \sup_{x \in [-\infty, \infty]^d} |(n - j)^{1/2} \hat{\gamma}_{j,ind}(x)|$$

$$= \max_{1+j \le t \le n} |(n-j)^{1/2} \hat{\gamma}_{j,ind}(Y_{t-j})|$$

The asymptotic quantile of KS(j) under the MDH can be approximated via a wild bootstrap approach. With the bootstrap critical values we can calculate uniform condence bands for $\hat{\gamma}_j(x)$ and the signicance of $\gamma_j(x)$ can be tested. The plot of a standardization of KS(j) against the lag parameter $j > 1$ can be viewed as generalization of the usual autocovariance plot in linear dependence to nonlinear conditional mean dependence. Escanciano and Velasco (2006b) called this plot the Integrated Pairwise Regression Functions (IPRF) plot.

3.5 Related hypotheses

In this chapter we have considered testing the MDH that, in statistical terms, just implies that the mean of an economic time series is independent of its past. The procedures studied in this chapter can be straightforwardly applied for testing the following generalization of the MDH

$$H_0 : E[Y_t|X_{t-1}, X_{t-2}, ...] = \mu$$

$\mu \in R$ where Y_t is a measurable real-valued transformation of X_t and $\mu = E[Y_t]$. This null hypothesis, which is referred to as the generalized MDH, contains many interesting testing problems as special cases. For instance, when Y_t is a power transformation of X_t, this null hypothesis implies constancy of conditional moments. The leading case in financial applications is the case where $Y_t = X_t^2$, because when X_t follows an MDS, this null hypothesis means that there is no volatility in the series X_t, that is, X_t is conditionally homoskedastic.

The cases $Y_t = X_t^3$ or $Y_t = X_t^4$ would respectively test for no dynamic structure in the third (conditionally constant skewness) and fourth (conditionally constant kurtosis), see for instance, Bollerslev (1987) and Engle and GonzÆlez-Rivera (1991). Another relevant case is when $Y_t = 1(X_t > c), c \in R^d$. In this case, the null hypothesis tested represents no directional predictability, see e.g. Linton and Whang (2007). Other situation of interest occurs when the null hypothesis of interest is the equality of the regression curves of two random variables, X_1t and X_2t, say; in this case, $Y_t = X_1t - X_2t$,$\mu = 0$.

Note also that most of the procedures considered in this chapter are also applicable for testing the null hypothesis that a general dynamic nonlinear model is correctly specified. In this situation, the null hypothesis of interest establishes that

$$\exists\theta_0 : E[\psi(Y_t, X_t, \theta_0)|X_t] = 0$$

where ψ is a given function, Y_t is a vector of endogenous variables and X_t is a vector of exogenous variables. Test statistics can be constructed along the lines described in this chapter. The main theoretical challenge in this framework is the way of handling the estimation of the parameters. There are basically three alternative approaches. First, to estimate the asymptotic null distribution of the relevant test statistics by estimating its spectral decomposition (e.g. Horowitz (2006) or Carrasco, Florens and Renault (2007)). Second, to use the bootstrap to estimate this distribution, see Wu (1986) and Stute, W., Gonzalez-Manteiga, W.G. and M. Presedo-Quindmil (1998). Third, to transform the test statistic via martingalization to yield an asymptotically distribution free test statistic.

Finally, in this chapter we have considered testing for MDS instead of testing for martingale. Recall that X_t is a martingale with respect to its natural filtration, when

$$E[X_t|X_{t-1}, X_{t-2}, ...] = X_{t-1}$$

Testing for martingale presents the additional challenge of handling nonstationary variables. Park and Whang (2005) considered testing that a first-order Markovian process follows a martingale by testing that the first difference of the process conditionally on the last value has zero mean, that is,

$$E(X_t - X_{t-1}|X_{t-1}) = 0$$

Hence, they allow for a singular nonstationary conditioning variable. This restrictive Markovian framework has the advantage of leading to tests statistics which are asymptotically distribution free, and hence, they do not need to transform their statistics or to use bootstrap procedures to obtain critical values.

Chapter 4

Conclusion

The previous chapters has presented a general panoramic on the literature of testing for the MDH. This area started at the beginning of the last century by developing tests for serial correlation and experimented a renewed interest recently because of the nonlinear dependence present in economic and, specially, financial series. The initial statistical tools were based on linear dependence measures such as autocorrelations or the spectral density function. These tools were initially considered motivated by the observation that economic time series follow normal distributions. Since in the last twenty five years it has been stressed the non normal behavior of financial series, the statistical and econometrics literature followed two alternative approaches. The first one targeted to robustify the well-established linear measures to allow for non-linear dependence. This approach has the advantage of its simplicity since it leads typically to standard asymptotic null distributions. However, its main limitation is that it cannot detect nonlinear dependence. The second approach considered nonlinear measures of dependence. Its advantage is that it is more powerful, its disadvantage is that asymptotic null distributions are nonstandard. Nowadays, this feature is hardly a drawback because the increasing availability of computing resources has allowed the implementation of bootstrap procedures that can estimate the asymptotic null distributions with relative ease.

The definition of martingale involves the information set of the agent that typically contains the infinite past of the economic series. This feature implies that, in practice, it is practically impossible to construct a test which, although it may be consistent theoretically, has power for any possible violation of the null hypothesis. The pairwise approach, which admittedly does not deliver consistent tests, leads to tests with reasonable power for common alternatives. Another sensible possibility to reduce this dimensionality problem is to consider alternatives of a single-index structure, i.e. where the conditioning set is given by a univariate, possibly unknown, projection of the infinite-dimensional information set. More research is clearly needed in this direction.

we have illustrate the different methodologies with exchange rate data that typically satisfy the MDH, as we have seen. Stock market data is not such a clear cut case. Rejecting the MDH leads to the challenge of selecting a proper model. In this respect, data-driven adaptive tests are informative, since they provide an alternative model in case of rejection. Notably, the principal component analysis provided in Escanciano and Mayoral (2007) represents a clear, theoretically well motivated approach, that coupled with an effective choice for the number of components can help in this selection process.

Andrews, D.W. K., and W. Ploberger (1996), Testing for Serial Correlation Against an ARMA (1, 1) Process,Journal of the American Statistical Association, 91, 1331 1342.

Barnett, W. and A. Serletis (2000), Martingales, nonlinearity, and chaos, Journal of Economics. Dynamics and Control, 24, 703-724

Bartlett, M.S. (1955), An Introduction to Stochastic Processes with Special Reference to Methods and Applications, Cambridge University Press, London.

Bekaert, G. and R. Hodrick (1992), Characterizing predictable components in excess returns on equity and foreign exchange markets, Journal of Finance, 467-509

Bera, A.K. and M.L. Higgins (1993) ARCH models: properties, estimation and testing, Journal of Economic Surveys, 7, 305-362.

Bera, A.K. and M.L. Higgins (1997) ARCH and bilinearity as competing models for nonlinear dependence, Journal of Business and Economic Statistics, 15, 43-51.

Beran, J. (1992), A Goodness-of-t Test for Time Series with Long Range Dependence, Journal of the Royal Statistical Society B, vol. 54, pp.749-760.

Bierens, H. J. (1982), Consistent model specication tests, Journal of Econometrics 20, 105-134.

Bierens, H. J. (1984), Model specication testing of time series regressions, Journal of Econometrics 26, 323-353.

Bierens, H. J. (1990), A consistent conditional moment test of functional form, Econometrica 58, 1443-1458.

Bierens, H. J. and W. Ploberger (1997), Asymptotic theory of integrated conditional moment tests, Econometrica 65, 1129-1151.

Bollerslev, T. (1987), A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, The Review of Economics and Statistics, 69, 542-547.

Bollerslev, T and J. M. Wooldridge, (1992) Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances,Econometric Reviews, 11, 143-179.

Box, G.E.P. and D. A. Pierce (1970) Distribution of residual autocorrelations in autoregressive integrated moving average time series models,Journal of the American Statistical Association 65, 1509-1526.

Brillinger, D.R. (1981), Time Series: Data Analysis and Theory, Holden Day, San Francisco.

Campbell, J. Y., A.W. Lo and A.C. MacKinlay (1997), The econometrics of nancial markets, Princeton University Press.

Carrasco, M., J.P. Florens, and E. Renault (2007), Linear Inverse Problems in Structural Econometrics, Handbook of Econometrics, Vol. 6, eds. J.J. Heckman and E.E. Leamer.

Chan, K.S. and H. Tong (2001), Chaos: a statistical perspective, Springer.

Chen, H. and J.P. Romano (1999), Bootstrap-assisted goodness-of-t- tests in the frequency domain, Journal of Time Series Analysis, 20, 619-654.

Clark,T.E. and K.D. West (2006), Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis, Journal of Econometrics, 135, 155186.

Cochrane, J. H. (1988), How big is the random walk in GNP?, Journal of Political Economy, 96 893-920.

Cochrane, J. H. (2005), Asset Pricing. Princeton University Press. Princeton, New Jersey.

Cumby, R. E. and J. Huizinga (1992) Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variable Regressions,Econometrica, 60, 185-196.

Dahlhaus, R. (1985), On the asymptotic distribution of Bartletts Up-statistic, Journal of Time Series Analysis, 6, 213-227.

Dahlhaus, R. and D. Janas (1996), A frequency domain bootstrap for ratio statistics in time series analysis, The Annals of Statistics, 24, 1934-1965.

de Jong, R. M. (1996), The Bierenstests under data dependence, Journal of Econometrics 72, 1-32.

Delgado, M. A., J. Hidalgo, and C. Velasco (2005), Distribution free goodness-of-t tests for linear models,The Annals of Statistics, 33, 2568-2609.

Delgado, M. A. and C. Velasco (2007), A new class of distribution-free tests for time series models specication, preprint.

Deo, R.S. (2000), Spectral tests of the martingale hypothesis under conditional heteroscedasticity, Journal of Econometrics, 99, 291-315.

Deo, R.S. and W. Chen (2000), On the integral of the squared periodogram, Stochastic Processes and their Applications, 85,159-176.

Diebold, F. X. (1986), Testing for serial correlation in the presence of heteroskedasticity, Proceedings of the American Statistical Association, Business and Economics Statistics Section, 323-328.

Diebold, F. X and J.A. Nason (1990), Nonparametric exchange rate predictions?, Journal of International Economics, 28, 315-332.

Dominguez, M. and I.N. Lobato (2003) Testing the martingale difference hypothesis, Econometric Reviews, 22, 351-377.

Durbin, J. and G.S. Watson, (1950), Testing for Serial Correlation in Least Squares Regression: I, Biometrika, 37, 409-428.

Durlauf, S. N. (1991), Spectral based testing of the martingale hypothesis, Journal of Econometrics, 50, 355-376.

Engle, R.F. and G. GonzÆlez-Rivera (1991), Semiparametric ARCH Models, Journal of Business & Economic Statistics, 9, 345-359.

Escanciano, J. C. (2007a), Model Checks Using Residual Marked Empirical Processes, Statistica Sinica, 17, 115-138.

Escanciano, J. C. (2007b): Weak convergence of non-stationary multivariate marked processes with applications to martingale testing, Journal of Multivariate Analysis, 98, 1321-1336.

Escanciano, J. C. (2008), On the Lack of Power of Omnibus Specication Tests, forthcoming in Econometric Theory.

Escanciano, J.C. and I.N. Lobato (2007), Data-Driven Portmanteau Tests for Testing for Serial Correlation, preprint

. Escanciano, J. C. and S. Mayoral (2007) Data-Driven Smooth Tests for the Martingale Difference Hypothesis, preprint.

Escanciano, J. C., and C. Velasco (2006a), Generalized Spectral Tests for the Martingale Difference Hypothesis,Journal of Econometrics, 134, 151-185.

Escanciano, J. C. and C. Velasco (2006b), Testing the Martingale Difference Hypothesis Using Integrated Regression Functions, Computational Statistics & Data Analysis, 51, 2278-2294.

Fama, E. F. (1991), Efficient Capital Markets: A Review of Theory and Empirical Work,Journal of Finance, 25, 383-417.

Fan, J. and L. Huang (2001), Goodness-of-Fit Tests for Parametric Regression Models,Journal of the American Statistical Association, 96, 640-652.

Ferreira, E., and W. Stute (2004) Testing for Differences Between Conditional

Means in a Time Series Context,Journal of the American Statistical Association, 99, 169-174.

Fong, W. M., S.K. Koh, and S. Ouliaris (1997), Joint variance ratio test of the martingale hypothesis for exchange rates, Journal of Business and Economics Statistics, 15, 51-59.

Fong, W. M. and S. Ouliaris (1995), Spectral tests of the martingale hypothesis for exchange rates, Journal of Applied Econometrics 10, 255-271.

Francq. C., R. Roy & J.-M. Zakoan (2005), Diagnostic checking in ARMA models with uncorrelated errors, Journal of the American Statistical Association, 13, 532-544.

Franke , J. and W. Hardle (1992), On bootstrapping kernel spectral estimates, The Annals of Statistics, 20, 121-145.

GonzÆlez, M and I.N. Lobato (2003), Contrastes de autocorrelacin, Gaceta de Economa, 14, pp. 41-57.

Grenander, U. and M. Rosenblatt (1957), Statistical Analysis of Stationary Time Series, Chelsea Publishing Company, New York.

Guay, A., and E. Guerre (2006), A Data-Driven Nonparametric Specication Test for Dynamic Regression Models,Econometric Theory, 22, 543-586.

Guerre, E., and P. Lavergne (2005), Rate-optimal data-driven specication testing for regression models, Annals of Statistics 33, 840-870.

Guo, B. and P. C.B. Phillips (2001), Testing for Autocorrelation and Unit Roots in the Presence of Conditional Heteroskedasticity of unknown Form,Cowles Foundation for Research in Economics, Yale University

Hall, R. E. (1978), Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence, Journal of Political Economy 86, 971-987.

Hart, J.D. (1997). Nonparametric Smoothing and Lack-of-Fit Tests. Springer Verlag, New-York.

Hinich, M. and D. Patterson (1992) A new diagnostic test of model inadequacy which uses the martingale difference criterion, Journal of Time Series Analysis 13, 233-252.

Hong, Y. (1996), Consistent Testing for Serial Correlation of Unknown Form, Econometrica, 64, 837-864.

Hong, Y., (1999), Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach, Journal of the American Statistical Association 84, 1201-1220.

Hong, Y. and T.H. Lee (2003), Inference on predictability of foreign exchange rate changes via generalized spectrum and nonlinear time series models, Review of Economics and Statistics 85, 1048-1062.

Hong, Y. and Y.J. Lee (2005), Generalized spectral tests for conditional mean models in time series with conditional heteroskedasticity of unknown form, Review of Economic Studies, 43, 499-541.

Horowitz, J. (2006) Testing a Parametric Model Against a Nonparametric Alternative with Identication Through Instrumental Variables, Econometrica, 74, 521-538.

Horowitz, J. L. and W. Hrdle (1994) Testing a parametric model against a semiparametric alternative, Econometric Theory 10, 821-848.

Horowitz, J., I.N. Lobato, J.Nankervis and N.E. Savin, (2006) Bootstrapping the Box-Pierce Q test: a robust test of uncorrelatedness, Journal of Econometrics, 133, 841-862.

Horowitz, J.L., and V.G. Spokoiny (2001) An adaptive, rate-optimal test of a parametric meanregression model against a nonparametric alternative, Econometrica 69, 599-632.

Hsieh, D. A. (1988), Statistical properties of daily foreign exchange rates, Journal of International Economics, 24, 129-145.

Hsieh, D. A. (1989), Testing for nonlinear dependence in daily foreign exchange rates, Journal of Business 62, 339-368.

Hsieh, D. A. (1993), Implication of nonlinear dynamics for nancial risk management, Journal of Finance and Quantitative Analysis, 28, 41-64.

Inglot, T., and T. Ledwina (2006), Data Driven Score Tests of Fit for Semiparametric Homoscedastic Linear Regression Model, Preprint 665, Institute of Mathematics Polish Academy of Sciences.

Inoue, A. and L. Kilian (2004), In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?, Econometric Reviews, 23(4), 371-402.

Janssen, A. (2000), Global power functions of goodness of t tests, Annals of Statistics, 28, 239-253.

Khmaladze, E. V. (1981), Martingale Approach in the Theory of Goodness-of-Fit Tests, Theory of Probability and its Applications 26, 240-257.

Koul, H. L. and W. Stute (1999), Nonparametric model checks for time series, The Annals of Statistics 27, 204-236.

Kuan, C.-M. and W. Lee (2004), A new test of the martingale difference hypothesis, Studies in Nonlinear Dynamics & Econometrics, 8.

Laudrup N. and M. Jansson (2006), Improving size and power in unit root testing, Palgrave Handbook of Econometrics, Vol. I, chapter 7, pp 252-277.

Lavergne, P. and V. Patilea (2007), One for all and all for one: dimension reduction for regression checks, forthcoming in Journal of Econometrics.

Lebaron, B. (1999), Technical trading rule protability and foreign exchange intervention, Journal of International Economics, 49, 125-143.

LeRoy, S. F. (1989), Efficient Capital Markets and Martingales,Journal of Economic Literature, 27, 39-44.

Levich, R. and L. Thomas (1993), The signicance of technical trading rule prots in the foreign exchange market: a bootstrap approach, Journal of International Money and Finance, 12, 451-474.

Linton, O. and Y.-J. Whang (2007), The Quantilogram: with an application to evaluating directional predictability, Journal of Econometrics, 141, 250-282.

Liu, R. Y. (1988), Bootstrap procedures under some non-i.i.d models, The Annals of Statistics 16, 1696-1708.

Liu, C. Y. and He, J. (1991), A variance ratio test for random walks in foreign exchange rates, Journal of Finance, 46, 773-785.

Ljung, G.M. and G. E. P. Box (1978), On a Measure of Lack of Fit in Time Series Models, Biometrika, 65, 297-303.

Lo, A. W. (1997), Market efficiency: stock market behaviour in theory and practice, Vol. I and II (Edward Elgar).

Lo, A.W. and A.C. MacKinlay (1989), The size and power of the variance ratio test in nite samples: a Monte Carlo investigation, Journal of Econometrics, 40, 203-238.

Lobato, I.N. (2001), Testing that a dependent process is uncorrelated, Journal of the American Statistical Association, vol.96, pp.1066-1076.

Lobato, I.N., J. C. Nankervis, and N.E. Savin (2001) Testing for autocorrelation using a modied Box-Pierce Q test,International Economic Review 42, 187-205.

Lobato, I.N., J. C. Nankervis, and N.E. Savin (2002), Testing for Zero Autocorrelation in the Presence of Statistical Dependence,Econometric Theory, 18, 730-743.

Lobato, I.N. and C. Velasco (2004), A General Test for White Noise, preprint.

McCurdy, T. H. and I.G. Morgan (1987), Testing of the martingale hypothesis for foreign currency futures with time-varying volatility, International Journal of Forecasting, 3, 131-148.

McCurdy, T. H. and I.G. Morgan (1988), Testing of the martingale hypothesis in Deutsche Mark futures with model specifying the form of heteroskedasticity, Journal of Applied Econometrics, 3, 187-202.

Meese, R. A. and K. Rogo (1983a), Empirical exchange rates models of the seventies, Journal of International Economics, 14, 3-24.

Meese, R. A. and K. Rogo (1983b), The out of sample failure of empirical exchange rates models: sampling error or misspecication? in J. Frenkel (Ed.), Exchange Rates and International Economics (Chicago: University of Chicago Press).

Milhłj, A. (1981), A test of t in time series models, Biometrika, 68, 177-187.

Nankervis J.C.,and N.E. Savin (2007) Testing for serial correlation: Generalized Andrews-Ploberger tests, preprint.

Neyman, J. (1937), Smooth Test for Goodness of Fit,Scandinavian Aktuarietidskr, 20, 149-199.

Paparoditis, E. (2000), Spectral density based goodness-of-t tests for time series models, Scandinavian Journal of Statistics 27, 143-176.

Park, J. Y., and Y.J. Whang (2005), Testing for the Martingale Hypothesis, Studies in Nonlinear Dynamics and Econometrics, Issue 2, Article 2.

Robinson, P. M.,(1991) Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regression,Journal of Econometrics, 47, 67-84.

Romano, J. L., and L.A. Thombs (1996), Inference for Autocorrelations Under Weak Assumptions, Journal of the American Statistical Association, 91, 590-600.

Shorack, G. and J. Wellner (1986), Empirical Processes with Applications to Statistics (Wiley: New York).

Stinchcombe, M. and H. White (1998), Consistent specication testing with nuisance parameters present only under the alternative, Econometric Theory 14, 295-325.

Stute, W. (1997), Nonparametric model checks for regression, The Annals of Statistics 25, 613-641.

Stute, W., W. G. GonzÆlez-Manteiga, and M. Presedo-Quindmil (1998), Bootstrap approximations in model checks for regression, Journal of the American Statistical Association, 83, 141-149.

Sweeny, R. J. (1986), Beating the foreign exchange market, Journal of Finance, 41, 163-182.

Weiss, A. (1986), ARCH and bilinear time series models: comparison and combination, Journal of Business and Economic Statistics, 4, 59-70.

Wooldridge, J. (1992), A test for functional form against nonparametric alternatives, Econometric Theory 8, 452-475.

Wu, C. F. J., (1986), Jacknife, bootstrap and other resampling methods in regression analysis with discussion, The Annals of Statistics 14, 1261-1350.

Yatchew, A. J. (1992) Nonparametric regression tests based on least squares, Econometric Theory 8, 435-451.

Yule, G.U. (1926), Why do we sometimes get nonsense correlations between time-series? A study in sampling and the nature of time-series, Journal of the Royal Society, vol. 89, pp.2-9, 30-41.

Zheng, X. (1996) A consistent test of functional form via nonparametric estimation technique, Journal of Econometrics 75, 263-289.

Stochastic Processes, (David Nualart) The University of Kansas nualart@math.ku.edu

F. C. Klebaner: (2003) Introduction to Stochastic Calculus with Applications.

http://maths.dur.ac.uk/stats/courses/StochProc34/

www.bath.ac.uk/ maspm/martingales.pdf

http://www.roulette.co.za/roulette-betting-systems/basics-martingale

www.math.uchicago.edu/ may/VIGRE/VIGRE2011/.../Ladd.pdf