CHAPTER 3

STOCK MARKET THEORY

3.1 INTRODUCTION

This chapter gives a detailed exposition of two theoretical models, the efficient market hypothesis and the present value model, which dominate the literature on stock market modeling. Intuitively the efficient market hypothesis asserts that stock prices adjust rapidly and unbiasedly to reflect new and relevant price sensitive information. Although these price adjustments are not always correct, it is unbiased so that the under- and over-adjustments are unpredictable. Since new information are random and independent and the large number of investors adjust stock prices rapidly to reflect this new information, price changes are independent and random (Reilly 1989:212). This means that trading based solely on historical prices, in other words technical analysis, cannot yield abnormal profits.

Initially, most of the empirical research on the efficient market hypothesis was done in terms of the random walk theory, which was based on empirical analysis without a theoretical foundation. Fama (1970) presented the first synthesis of the efficient market theory in terms of the fair game model and Samuelson (1965) and Mandelbrot (1963) showed that the fair game model is analytically equivalent to the expected present value theory of security valuation (also called the present value model). According to this theory, stock prices are a function of all the expected future dividends discounted at the discount rate.

Although the present value model provides a theoretical foundation for modeling stock prices, the macroeconomic variables that can be used to estimate it empirically has to be identified since the explanatory variables in the model is in practice usually replaced by proxies and supplemented by additional variables. Therefore, this chapter extends the theoretical foundation with some empirical implications.

Recently, the potential asymmetry in stock market behavior has received considerable attention in the literature. It is generally assumed in investment theory that investors are risk averse, since any rational investor will prefer certainty to risk *ceteris paribus* (Reilly 1989:10,255; Renwick 1971:400). This risk aversion leads to asymmetric behavior on the part of investors, which in turn causes asymmetry in the behavior of their investments, especially in the stock market. In section 3.3 an overview of the theory underlying stock market asymmetry is presented.

3.2 THE EFFICIENT MARKET HYPOTHESIS AND THE PRESENT VALUE MODEL

3.2.1 The Efficient Market Hypothesis and Implications for Stock Market Modeling

The efficient market hypothesis holds that prices adjust rapidly and unbiasedly to new and relevant price sensitive information. The three assumptions underlying the efficient market hypothesis are: (i) A large number of profit-maximizing investors that operate independently of each other. (ii) New information regarding securities comes to the market in a random fashion and the announcements over time are generally independent of one another. (iii) Investors adjust security prices rapidly to reflect the effect of the new information. While these price adjustments are not always correct, it is unbiased so that the under- and over-adjustments are unpredictable. Since new information are random and independent and the large number of investors adjust stock prices rapidly to reflect this new information, price changes are independent and random (Reilly 1989:212).

Three forms of the efficient market hypothesis exist namely the weak form, semi-strong form and the strong form (Marx *et al* 2003:35). According to the strong form of the efficient market hypothesis, security prices fully reflect all the relevant public and private information and that financial markets are "perfect markets" in the sense that all information is free and available to everybody simultaneously. The semi-strong form relaxes these assumptions somewhat and assumes that security prices adjust rapidly to reflect all public information, which are defined as market information as well information such as economic and political news and company news such as earnings, dividend announcements, mergers and acquisitions. The weak form assumes that security prices adjust rapidly to reflect all security market information, including security prices, trading volume and rates of return.

The efficient market hypothesis has crucial implications for stock market investors and their approach to stock market trading. If capital markets are efficient and security prices fully reflect all relevant information as postulated by the efficient market hypothesis, any trading rule solely based on past data cannot yield above-average returns, since by the time the information is public, the price adjustment has taken place. Therefore, trading based on technical analysis where the basic philosophy is that security prices tend to move in trends so that their trading rules are past price movements will not be able to yield above-average returns (Reilly 1989:658, 245).

In contrast, the philosophy underlying fundamental analysis is that the intrinsic value of a security is determined by the underlying economic variables. Fundamental analysts analyze and estimate macroeconomic prospects such as economic growth, inflation and interest rates and then identify industries that will gain most from these conditions. The fundamental analyst subsequently determine the intrinsic value of the companies within these industries, in order to identify and invest in those that are undervalued, that is, for which the market price is lower than the intrinsic value. It is possible for the market price and the intrinsic value of a security to differ, but the market will eventually correct this discrepancy. Therefore, an analyst capable of making a better than average estimate of the intrinsic value will be able to make above-average profits.

Initially, most of the empirical research on the efficient market hypothesis was done in terms of the random walk theory, which was based on empirical analysis without a theoretical foundation. Fama (1970) presented the first synthesis of the efficient market theory in terms of the fair game model. In contrast to the random walk hypothesis, which dealt with the behaviour of prices over time, the fair game model focuses on the price in a specified period. The model is built on the assumption that the security's price fully reflects all the relevant information available up to that period. A fair game model is defined as follows: Let the expected theory of price formation be described by the following equation:

$$E(P_{j,t+1} \mid \phi_t) = [1 + E(P_{j,t+1} \mid \phi_t)]P_{j,t} \tag{3.1}$$

where E is the expected value operator, $P_{j,t}$ is the price of security j in period t, $r_{j,t}$ is the rate of return on security j during period t and ϕ_t is the shared information set. Further, let $x_{j,t+1}$ be the difference between the actual and expected prices in period t:

$$x_{j,t+1} = P_{j,t+1} - E(P_{j,t+1} \mid \phi_t)\,. \tag{3.2}$$

The sequence $\{x_{j,t}\}$ is called a "fair game" if $E(x_{j,t} \mid \phi_t) = 0$. According to the efficient market hypothesis this should be the case for $x_{j,t}$ since it is impossible to consistently earn abnormal returns based on the shared information due to the competition between investors.

3.2.2 The Present Value Model

Samuelson (1965) and Mandelbrot (1963) showed that the fair game model (see section 3.2.1) is analytically equivalent to the expected present value theory of security valuation (also called the present value model). Assume that return on any security i in period t, $r_{i,t}$, minus a security specific constant k_i is a fair game, which means that

$$E(r_{j,t} \mid \phi_t) - k_i = 0 . \tag{3.3}$$

By substituting the definition of returns as the dividend yield ($D_{i,t-1}/P_{i,t}$) plus capital gains ($(P_{i,t-1}-P_{i,t})/P_{i,t}$), the equation becomes

$$\frac{E[(P_{i,t-1} + D_{i,t-1}) \mid \phi_t] - P_{i,t}}{P_{i,t}} - k_i = 0 . \tag{3.4}$$

Which is equivalent to

$$P_{i,t} = \frac{E[(P_{i,t-1} + D_{i,t-1}) \mid \phi_t]}{(1 + k_i)} . \tag{3.5}$$

By substituting for P_{t+1} and replacing t with t+1, equation 3.5 becomes

$$P_{i,t} = \frac{E[(D_{i,t}) \mid \phi_t]}{(1 + k_i)} + \frac{E[P_{i,t+2} + D_{i,t+2}]}{(1 + k_i)^2} . \tag{3.6}$$

If it is assumed that $k_i>0$ and this process is repeated n times, then

$$\lim_{n \to \infty} \frac{E[D_{i,t+n}]}{(1 + k_i)^n} = 0 . \tag{3.7}$$

This yields the familiar expected present value model first presented by Smith (1925) and Burr-Williams (1938):

$$P_{i,t} = \sum_{n=1}^{\infty} \frac{E[D_{i,t+n}]}{(1 + k_i)^n} . \tag{3.8}$$

The current share price can be solved from equation 3.8 by setting t=0

$$P_{i,0} = \sum_{n=1}^{\infty} \frac{E[D_{i,n}]}{(1+k_i)^n}. \tag{3.9}$$

Equation 3.9 shows that the price of a security is equal to the present value of the expected future dividend receipts of the asset. In this formula the expected capital gain realized upon the sale of the security is subsumed, since its magnitude also depends on the present value of the expected future dividend stream. Under the assumption that expected dividends grow at a constant rate, Gordon and Shapiro (1956) replaced the problem of forecasting an infinite number of future dividends with that of estimating a single expected growth rate g. This means that equation 3.9 can be written as

$$P_{i,0} = \sum_{n=1}^{\infty} \frac{D_{i,0}(1+g_i)^n}{(1+k_i)^n} \tag{3.10}$$

where $P_{i,0}$ is the price of security i in period 0, $D_{i,0}$ is the dividend in period 0, g_i is the expected growth rate of security i and k_i is the rate at which the dividends are discounted. By using the properties of the sum to infinity of a geometric series, equation 3.10 can be reduced to the *constant growth model*

$$P_{i,0} = \frac{E(D_{i,1})}{k_i - g_i} = \frac{D_{i,0}(1+g_i)}{k_i - g_i}. \tag{3.11}$$

Therefore, the equilibrium prices of security i is determined by its dividend (D_i), the growth rate (g_i) and the discount rate (k_i).

3.3 EMPIRICAL IMPLICATIONS OF THE PRESENT VALUE MODEL

In practice the explanatory variables in the present value model (dividends, growth and the discount rate) is in practice usually replaced by proxies and supplemented by additional variables. Therefore, although the present value model provides a theoretical foundation for modeling stock prices, the macroeconomic variables that can be used to estimated it empirically still has to be identified.

3.3.1 The Discount Rate

In order for the present value model to be useful for empirical analysis, the discount rate (k_i) has to be defined more specifically. The discount rate is determined by three factors: (i) the economy's real risk-free rate, (ii) the expected rate of inflation and (iii) a risk premium (Reilly 1989:326). Investors want to be compensated for expected inflation, so that their money does not loose purchasing power over time. In addition, they want to receive the real risk-free rate to compensate them for the opportunity cost of parting with their money. Finally, a risk premium is added to the discount rate to compensate for the uncertainty regarding the expected returns of the security.

(i) The risk premium

Unlike the expected inflation and real risk-free rate, the risk premium for different assets may differ, reflecting the different risk or uncertainty of their returns. A risk-free investment can be defined as an investment of which both the amount and timing of the expected income stream are certain. However, the timing and amount of expected income from most investments are not certain and hence investors require a risk premium on top of the risk-free rate to compensate for the risk involved in their investment. The risk associated with investment includes several major sources of uncertainty, namely business risk, liquidity risk, exchange rate risk, interest rate risk, purchasing power risk, management risk, default risk and industry risk (Marx *et al* 2003:174). These risks can be categorized as either firm-specific risks, which differs between the securities of different firms, and general risks, which are common across all the securities of firms operating in a particular country.

The firm-specific risks associated with a security can broadly be divided into industry risk, business risk and management risk. Industry risk includes the uncertainty of operating within a particular industry, such as the financial services or natural resources industries. This risk influences all the firms operating in this particular industry. The uncertainty of income flows caused by the nature of the firm's business is known as business risk (Reilly 1989:16). Usually investors consider the distribution of a firm's income and assign a risk premium accordingly. The uncertainty of income caused by the basic business of the firm is typically measured by the distribution of the firm's operating income (defined as earnings before interest and taxes) over time. The more volatile the firm's operating income is over time relative to its mean income, the greater the business risk. Finally, management risk is the uncertainty or risk introduced by the management team of the specific firm in terms of their management style and strategy.

Related to the firm-specific risks mentioned above, securities are also subjected to liquidity and default risk which are both to a large extent firm-specific. Default risk refers to the variability of returns caused by changes in the creditworthiness of the firm in which the investment is made (Marx *et al* 2003:10). Liquidity risk is the risk of not being able to quickly convert an asset into cash without a substantial price concession (Reilly 1989:16). The greater the uncertainty of being able to sell an asset quickly without a loss, the greater the liquidity risk.

In addition to these firm-specific risks, securities are also subject to the risks of investing in the particular country in which the firm operates, which are homogenous across investments within that country. Interest rate risk is the potential influence of changes in the market interest rate on returns (Marx *et al* 2003:9). The value of a security is determined by discounting all future income expected from the security to determine its present value and therefore the value will move inversely with changes in market interest rates. The uncertainty regarding the future behavior of interest rates therefore poses a risk in the sense that it introduces uncertainty regarding the correct valuation of share prices and hence a possible loss (or gain) to the extent that this valuation differs from the true intrinsic value of the share(s).

Purchasing power or inflation risk is the uncertainty or risk associated with changes in the inflation rate. The returns generated by the investment should be sufficient to at least keep pace with inflation and hence preserve the purchasing power of the initial investment. Fluctuations in the rate of inflation therefore introduces an uncertainty regarding the required rate of return as well as the risk that the purchasing power of the investment will deteriorate due to future increases in inflation.

Exchange rate risk is the uncertainty involved when investing in a foreign currency (Reilly 1989:16). Globalization and deregulation have made a portfolio comprising international assets universal. When investing globally, the return on an investment in the foreign currency has to be converted to the domestic currency to calculate the return for the investor. This means that investors have to take into account the risk that the exchange rate between their domestic currency and the foreign currency in which the investment is made might change.

Although the risks discussed in the previous paragraphs are relevant for the pricing of a particular security, the firm-specific risks became redundant in the pricing of the general stock market. Since the aggregate stock market reflects the average of the prices of individual shares, only the risks influencing all shares such as interest rate risk and purchasing power risk are relevant in pricing the aggregate stock market. The firm-specific risks are then added to this when a specific share is priced. Therefore the risk premium used in modeling the stock market should reflect only the risks that are not firm-specific.

(ii) Interest rates

According to the expectations hypothesis, the shape of the term structure is explained by the interest rate expectations of market participants. The long-term interest rate is the rate that the long-term investor would expect to earn through successive investments in short-term securities over an investment horizon equal to the term to maturity of the long-term issue. In other words, the long-term interest rate on a security is the average of all the expected future short-term interest rates expected to prevail over the duration of the security. Expressed algebraically:

$$(1+{}_tR_n)=[(1+{}_tR_1)({}_{1+t+1}r_1)\ldots({}_{1+t+n-1}r_1)]^{1/N}$$

where R_n is the actual long-term interest rate, N is the term to maturity (in years) of the long-term issue, R is the current one-year rate and ${}_{t+1}r_1$ is the expected one-year yield during some future period t+1 (Reilly 1989:416).

According to the expectations hypothesis if market participants expect short-term interest rates to rise in the future then the long-term interest rate will be higher than current short-term interest rates and the yield curve will be upward-sloping. On the other hand, if short-term interest rates are expected to fall then the long-term interest rate will be below short-term interest rates, so that the yield curve will be downward-sloping.

Since the expected future short-term interest rates are used to discount future returns in a present value model of stock prices, the long-term rate can be used as discount factor since it captures the expected short-term interest rates. Harasty and Roulet (2000) showed that the stock market is cointegrated with the long-term interest rate and a proxy for dividends in 17 developed countries. Zhou (1996) interest rates significant in explaining stock returns. Ansotegui and Esteban (2002) showed that the stock market is cointegrated with industrial production, inflation and the long-term interest rate.

3.3.2 Dividends and Growth

Theoretically, the present value model asserts that security prices are determined by dividends and the discount rate. It follows trivially that any factor that influences the stream of cash flows or the discount rate will systematically influence stock prices. Since the seminal article by Chen *et al* (1986), the influence of variables such as interest rates and inflation on the discount rate and of the industrial production growth on the expected cash flows or dividends has been well established. In empirical studies, these and other variables are usually used to proxy the influence of dividends on stock prices. Jondeau and Nicolai (1993) have shown that only in the US do dividends directly explain stock prices and in other countries dividends have to be replaced by proxies.

In the literature, dividends are usually replaced by proxies such as industrial production, unemployment or the state of the business cycle in empirical analyses. Since most firms are more profitable during periods of high economic growth than low economic growth, the state of the business cycle should positively impact on the stock market through its influence on dividends. Brocato and Steed (1998) have shown that total returns of equity assets rise (fall) during expansions (recessions). Ansotegui and Esteban (2002) used industrial production as proxy for dividends and showed that the stock market is cointegrated with industrial production, inflation and the interest rate. In contrast, Domian and Louton (1995 and 1997) showed that the unemployment rate, which is often used as an indicator of the state of the economy, influences the stock market returns in the US. Since the stock market tends to be forward-looking, several studies have modeled stock market returns as a function of the future (instead of current) state of the business cycle. For example, Chen (1991) has shown that variables reflecting the future state of the economy are positively related to the expected excess market return.

Fama and Schwert (1977), Solnik (1983) and Spyrou (2001) found a negative relationship between inflation and stock market returns and ascribed this to the negative correlation between inflation and real output growth. In other words, since stock returns are positively related to real activity and real activity is negatively related to changes in the level of prices, stock returns are negatively related to inflation. Kaul (1990) empirically proves that this is also the case for the US stock market. Ansotegui and Esteban (2002) proxied dividends with industrial production and showed that the stock market is cointegrated with inflation, industrial production and the interest rate.

3.4 STOCK MARKET ASYMMETRY

It is generally assumed in investment theory that investors are risk averse (Reilly 1989:10, 255). When an investor is faced with a choice between two investments with the same expected rate of return, he/she will choose the one with the smallest risk. Theoretically, investors are risk averse since their utility functions are assumed to exhibit decreasing marginal utility of wealth. In other words, they have concave utility functions, or mathematically

$$U''(w) < 0 \tag{3.12}$$

where w is wealth, U is utility and U″ is the second derivative of the utility function with respect to wealth.

Rabin and Thaler (2001) argued that an explanation for risk aversion should also incorporate loss aversion. Loss aversion refers to the inclination of economic agents to be more sensitive to reductions in their levels of well-being than to increases (Benartzi and Thaler 1995). Kahneman and Tversky (1979) presented the following loss aversion utility function:

$$-\lambda E\left[(-\chi)^{(1-\gamma_1)}\mathbf{1}_{\{\chi\leq 0\}}\right] + E\left[(-\chi)^{(1-\gamma 2)}\mathbf{1}_{\{\chi>0\}}\right] \tag{3.13}$$

where **1** is an indicator variable, $\chi = W - B_0$ is the gain or loss of final wealth (W) relative to a benchmark B_0. This utility function has the problem that it gives different preferences if χ is expressed in different units unless γ_1=γ_2 or the difference between γ_1 and γ_2 is small. However, this can be overcome by expressing χ in returns.

Risk and loss aversion mean that investors are more sensitive to losses than gains, which means that will behave differently when expecting gains than when expecting losses. In other words, their behavior is asymmetric with respect to positive or negative returns.

Therefore, since the stock market is driven by the behavior of investors, this potentially causes asymmetry in stock prices as well.

Two explanations have been given in the literature on why investors' risk and/or loss aversion induces stock market asymmetry. First, Chalkley and Lee (1998) argues that risk aversion encourages economic agents to react promptly on receiving bad news, while it prevents them from acting quickly when receiving good news. The cautious response of investors at the individual level aggregates to result in economy-wide asymmetric behavior. They developed a complex Markov-switching model showing that the risk aversion of economic agents does result in asymmetric behavior conditional on the state of the business cycle. However, the derivation of their model is beyond the scope of this study and only their intuitive explanation will be presented here.

Assume that investors are faced with a choice between high activity and low activity and that the desirability of each choice depends on the state of nature, so that low activity is preferred when state is bad and vice versa. Suppose also that low activity is safer, so that the expected utility of low activity is higher when both states are equally likely. Then investors will require a strong belief in favor of the good state to choose the high level of activity, but only a weak belief in the bad state to choose the low level of activity.

A downturn in the economic data that is used to evaluate share prices may be indicative of other economic agents receiving bad news or it might be a random change, but in either case the cautious response is to choose low activity. In this case (a downturn in economic data), risk aversion and uncertainty about the information value of aggregate data work together, leading informed agents to quickly respond to bad news or conditions and other agents to quickly respond to that response. Of course, there is also uncertainty about the interpretation of an upturn in the relevant economic data, but in this case risk aversion works against reacting to such a signal. When agents are reluctant to react to an upturn in economic data signals, it is reasonable to infer that an upturn is more likely a consequence of noise than any genuine good news and hence reticence is rational.

Therefore, it can be expected that investors will react more reluctantly to good news, expectations and vice versa. When the behavior of these individual investors are aggregated it implies that the stock market will react quicker during good conditions or on good news or expectations, or put differently, that its adjustment to equilibrium will be slower during adverse economic conditions and faster during positive economic conditions. The "upturn" and "downturn" of data in the Chalkley and Lee (1998) framework originally referred to good or bad conditions as reflected in the business cycle. According to the present value model stock prices are a function of the discount rate and dividends and since real economic activity is one of the main determinants of dividends an economic upswing (downswing) will cause higher (lower) dividends and can therefore be considered as good (bad) news or conditions.

The second explanation for asymmetric investor (and hence stock market) behavior is driven by the potential loss (profit) in an overvalued (undervalued) stock market. Following the same line of reasoning as Chalkley and Lee (1998), Phelps and Zoega (2001) and Siklos (2002) also hypothesized different speeds of adjustment but they introduced a different driving force for the asymmetry by redefining the good and bad news or conditions that prompts the asymmetric behavior of investors.

Their theory on stock market asymmetry is based on the paradigm of the structural slump developed by Phelps (1967). A structural slump is characterized by a steep decline in share prices followed by a gradual rise in unemployment. A structural boom, on the other hand, entails a steep rise in share prices followed by a decline in unemployment. In the case of a structural boom, investors calculate that this signals a jump in future asset returns and, consequently, the valuation of these assets as reflected in the stock market. The resulting rise in the profitability of investment signals a falling unemployment rate. The boom ends when the productivity rise increases investment costs.

Theoretically, this scenario works symmetrically, but Phelps and Zoega (2001) argued that it might in practice work asymmetrically since other factors may influence the progress of the business cycle. The potential asymmetry was first tested empirically by

Siklos (2002). His results showed that the relationships between the economy and the stock markets of the UK and the US were indeed asymmetric.

Although Siklos (2002) tested the stock market asymmetry based on the relationship between the stock market and unemployment, the asymmetry also holds for any other stock market model. If the stock market is undervalued it means that the market prices of shares are below their intrinsic value, so that a profit opportunity created since investors can buy shares at the low current market price and eventually resell it at a higher price once the market has corrected the discrepancy between the market and intrinsic value. In contrast, when the stock market is overvalued market prices of shares are above the intrinsic values. Eventually the market will correct this discrepancy so that share prices fall, in which case investors will loose money. Since investors are loss averse it is more important to avoid the potential loss if the market is overvalued than to make the profit if the market is undervalued. They will therefore react faster to an overvaluation that poses a potential loss than to an undervaluation that poses a potential profit.

3.5 CONCLUSION

This chapter reviewed the basic theoretical foundations, the efficient market hypothesis and the present value model for modeling stock markets. According to the efficient market hypothesis, capital markets are efficient in the sense that stock prices adjust rapidly and unbiasedly to reflect new and relevant price sensitive information. The price adjustment, although not always correct, is unbiased so that the under- and over-adjustments are unpredictable. Since new information are random and independent and the large number of investors adjust stock prices rapidly to reflect this new information, price changes are independent and random (Reilly 1989:212). This means that trading based solely on historical prices, in other words technical analysis, cannot yield abnormal profits.

Fama (1970) presented the first synthesis of the efficient market theory in terms of the fair game model and Samuelson (1965) and Mandelbrot (1963) showed that the fair game model is analytically equivalent to the expected present value theory of security valuation. According to this theory, stock prices are a function of all the expected future dividends discounted at the discount rate. Dividends are usually replaced in empirical studies with proxies measuring the state of the business cycle or the performance of the aggregate economy. The discount rate can be constructed as the sum of the risk-free rate and a risk premium.

Rational investors are assumed to be risk and loss averse and this potentially leads to asymmetric behavior that in turn results in stock market asymmetry. Two explanations have been given in the literature on why this might cause asymmetric investment behavior. First, it encourages investors to react promptly on receiving bad news while it prevents them from acting quickly when receiving good news (Chalkley and Lee 1998). Second, it is more important to avoid the potential loss if the market is overvalued than to profit if the market is undervalued and hence investors will react faster in an overvalued market. The cautious response of investors at the individual level aggregates to result in economy-wide asymmetric behavior.

Therefore, any empirical stock market model has to be build on the theoretical foundation of the present value model, taking into account the implications of the efficient market hypothesis as well as the potential asymmetry caused by investors' risk and loss aversion. In the next chapter studies modeling stock markets will be reviewed in order to analyze the empirical implications of the present value model, the efficient market hypothesis and investor asymmetry.