



## **Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin**

### **Copyright statement**

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

### **Liability statement**

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

### **Access Agreement**

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.



**EQUITY PRICING, VOLATILITY AND SKEWNESS**  
**The Dynamic Behaviour and Interactions of First and Second Moments in the**  
**Euro Area and United States Stock Markets**

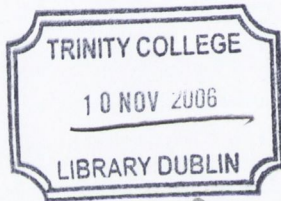
**Valerio Poti**

A dissertation submitted in fulfilment of the requirements for the Degree of Doctor  
of Philosophy in Business Studies

Supervisor: Prof. Colm Kearney

School of Business Studies  
Trinity College, Dublin

3<sup>rd</sup> August 2006



THO 81 S  
7991

*I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy in Business Studies, has not been submitted as an exercise for a degree at this or any other University, is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work. I also agree that the Library may lend or copy the thesis upon request.*

Signed: Valerio Bot.

ID No.: 02163290

Date: 23/08/2006



## Summary

This thesis contributes to the empirical asset pricing literature on both the cross-section and time series of stock returns. It also contributes to the recent but rapidly growing literature on total and idiosyncratic risk and to the literature on the dependence structure of stock market returns.

Concerning the asset pricing problem, the main contribution is a novel beta-pricing representation of Harvey and Siddique's (2000) 3-moment conditional CAPM. Its main advantage is that both its beta coefficients and risk premia can be interpreted as parameters of appropriate regression models. In an empirical application to US stocks sorted into 30 industry portfolios, I add to the extant evidence that, while coskewness helps explain a substantial portion of the cross section of average returns and the coskewness premium is of the same order of magnitude as the covariance premium, the estimated unconstrained 3-moment model implies a utility function that is not concave over the relevant wealth range.

To study the behaviour over time of total and idiosyncratic risk, I construct a unique and comprehensive firm and industry-level dataset of stock returns and volatilities in the Euro area, I derive a version of Campbell, Lettau, Malkiel and Xu's (2001) decomposition of total risk into idiosyncratic and systematic risk based on returns instead of excess returns and I define an original average correlation index. This average correlation index is very useful to study the aggregate dynamic behaviour of the correlations between a large number of assets. I find that European stocks have indeed become more volatile, and that idiosyncratic risk is the largest component of this volatility. I also find that the potential benefit of diversification strategies in Europe remains substantial and relatively stable over time.

I confirm earlier findings based on US data, e.g. Whitelaw (1994) and Goyal and Santa Clara (2003), that the simple lagged market variance-return relation is positive

but statistically insignificant. I question, however, Goyal and Santa Clara's (2003) conclusion that the return on the market portfolio is positively related to the idiosyncratic component of lagged total volatility. Instead, I find that when market volatility is also included in a long-run predictive regression, the relation between market return and average idiosyncratic volatility is negative. This is a striking confirmation, obtained using Euro area data, of a similar finding recently reported by Guo and Savickas (2003) for a portfolio of the 500 largest US stocks. I show that this result depends on the circumstance that market and average idiosyncratic volatility jointly proxy for average correlation and thus, perhaps surprisingly, for a component of systematic risk.

I also estimate the correlations of Euro-area market, industry and sector stock indices in a large-scale multivariate conditional setting, using extensions of Engle's (2002) Dynamic Conditional Correlation GARCH model. For comparison, I also apply the same estimation strategy to a sample of Euro area stock with largest market capitalization. I do not find evidence of a deterministic time trend either in individual stock and industry correlations or in sector or country level equity correlations. The latter however dramatically increase in recent years due to a structural break shortly before the launch of the Euro. These findings imply that, while industry and firm level diversification have retained their effectiveness, investment managers engaged in asset allocation in the Euro area should not rely any longer on country level diversification strategies.

Finally, I derive original analytical results on the relation between discount factor volatility and conditional return volatility. This relation provides the amount of discount factor volatility that rational asset pricing models must generate in order to explain conditional return volatility. In an empirical application to the US stock market, I find that it takes about 7 percent discount factor volatility to explain the conditional return volatility from 1871 to 2003, given about 4 percent dividend growth volatility.



## **Acknowledgements**

I would like to thank my Supervisor at Trinity College, Prof. Colm Kearney for his endless academic help and support. He introduced me to academic research on financial market volatility. I would like to thank him for his precious guidance and advice at every stage of my work.

I also would like to thank, more or less in order of appearance on the scene of my life, my mother, my father, my wife and my two children. This thesis would not exist without my mother's determination, throughout my young years, to let me avail of nothing less than the best educational opportunities and without her vision and enthusiasm for life. This thesis would not exist without my father's words that, at various stages of my life, have inspired me and given me strength, reminding me to keep in touch with real world practicalities while trying not to get stuck with them. This thesis, like many other nice things that have happened since I have met my wife, would not exist without Maria's support and her contagious capacity to dream. This thesis also owes a lot to my two children, who have filled my heart with joy and enthusiasm. Matteo's running around the house and Marco's kicking in his mum's belly are truly the best boost to creativity that I could ask for.



## TABLE OF CONTENTS

<b>CHAPTER 1: INTRODUCTION AND OVERVIEW .....</b>	<b>9</b>
1.1. INTRODUCTION.....	9
1.2. CONDITIONAL VS. UNCONDITIONAL MOMENTS .....	9
1.3. OLD AND NEW PARADIGM .....	12
1.4. THE FUNDAMENTAL RESEARCH QUESTIONS .....	14
1.5. MOTIVATIONS .....	16
1.6. STRUCTURE OF THE THESIS .....	19
1.7. MAIN FINDINGS AND CONTRIBUTIONS .....	20
1.8. SUMMARY AND CONCLUSIONS .....	23
<b>CHAPTER 2: THE ASSET PRICING LITERATURE .....</b>	<b>24</b>
2.1. INTRODUCTION.....	24
2.2. STOCHASTIC DISCOUNT FACTOR PRICING.....	25
2.2.1 <i>Alternative Representations</i> .....	25
2.2.2 <i>Excess Returns and the Mean of the SDF</i> .....	28
2.3. FACTOR MODELS .....	28
2.4. ASSET PRICING PARADIGMS AND EMH .....	33
2.5. RATIONAL ASSET PRICING MODELS .....	34
2.5.1 <i>Consumption Asset Pricing</i> .....	34
2.5.2 <i>Representative Investor</i> .....	36
2.5.3 <i>CAPM</i> .....	38
2.5.4 <i>(C)CAPM</i> .....	39
2.5.5 <i>Conditioning Variables</i> .....	41
2.5.6 <i>The Role of Systematic Skewness</i> .....	49
2.5.7 <i>The 3M-CAPM</i> .....	49
2.5.8 <i>The 3M-(C)CAPM</i> .....	52
2.5.9 <i>Tests of the 3M-(C)CAPM</i> .....	54
2.5.10 <i>3M-CAPM vs. (C)CAPM</i> .....	55
2.5.11 <i>The Role of Idiosyncratic Income Risk</i> .....	56
2.5.12 <i>The Market Risk Premium in Integrated and Segmented Markets</i> ...	58
2.6. BEHAVIOURAL MODELS.....	61
2.7. ABSOLUTE VS. RELATIVE PRICING .....	63
2.8. SUMMARY AND CONCLUSIONS.....	64
<b>CHAPTER 3: THE SECOND MOMENTS LITERATURE .....</b>	<b>66</b>
3.1. INTRODUCTION.....	66
3.2. SECOND MOMENTS .....	67
3.2.1 <i>Systematic vs. Idiosyncratic Volatility</i> .....	67
3.2.2 <i>Clustering</i> .....	69
3.2.3 <i>Asymmetry</i> .....	69
3.2.4 <i>Skewness</i> .....	73
3.2.5 <i>Contagion</i> .....	76
3.3. VARIANCE DECOMPOSITION.....	78



3.4.	UNCONDITIONAL AND CONDITIONAL SECOND MOMENTS ESTIMATES .....	82
3.4.1	<i>Unconditional Estimates</i> .....	82
3.4.2	<i>Conditional Estimates</i> .....	86
3.4.3	<i>DCC-GARCH</i> .....	107
3.4.4	<i>DCC-GARCH Extensions</i> .....	110
3.4.5	<i>Financial Applications</i> .....	111
3.5.	SUMMARY AND CONCLUSION .....	117
<b>CHAPTER 4: THE CROSS-SECTION OF STOCK RETURNS .....</b>		<b>119</b>
4.1.	INTRODUCTION .....	119
4.2.	THE 3M-(C)CAPM BETA PRICING REPRESENTATION .....	120
4.3.	DATA .....	122
4.4.	BETA-PRICING EMPIRICAL SPECIFICATION .....	122
4.5.	CONCAVITY OF THE UTILITY FUNCTION AND THE 3M-CAPM .....	129
4.6.	THE 3M-CAPM VS. THE (C)CAPM .....	131
4.7.	GMM ESTIMATES WITH RA AND NIARA .....	134
4.8.	SUMMARY AND CONCLUSIONS .....	138
<b>CHAPTER 5: EUROPEAN EQUITY RETURNS SECOND MOMENTS .....</b>		<b>140</b>
5.1.	INTRODUCTION .....	140
5.2.	VARIANCE DECOMPOSITION .....	141
5.3.	EMU DATA AND VARIABLE CONSTRUCTION .....	146
5.4.	TIME SERIES BEHAVIOR OF UNCONDITIONAL SECOND MOMENTS .....	151
5.5.	US VS. EMU SERIES .....	161
5.6.	WHAT MIGHT EXPLAIN VOLATILITY AND $R^2$ TRENDS? .....	161
5.7.	SECOND VS. OTHER MOMENTS .....	166
5.8.	SUMMARY AND CONCLUSIONS .....	175
<b>CHAPTER 6: CORRELATIONS OF EUROPEAN EQUITY RETURNS ...</b>		<b>176</b>
6.1.	INTRODUCTION .....	176
6.2.	DATA AND SUMMARY STATISTICS .....	177
6.3.	EMU COUNTRY AND STOCK CORRELATIONS .....	182
6.3.1	<i>Unit Root Tests</i> .....	184
6.3.2	<i>Wald-Type Tests</i> .....	187
6.3.3	<i>DCC-MVGARCH</i> .....	188
6.4.	SUMMARY AND CONCLUSIONS .....	201
<b>CHAPTER 7: THE TIME SERIES OF STOCK MARKET RETURNS .....</b>		<b>203</b>
7.1.	INTRODUCTION .....	203
7.2.	DISCOUNT FACTOR AND CONDITIONAL RETURN VOLATILITY .....	203
7.3.	THE MODEL .....	204
7.4.	EMPIRICAL APPLICATION .....	208
7.5.	SUMMARY AND CONCLUSION .....	213
<b>CHAPTER 8: IMPLICATIONS, LIMITATIONS AND FUTURE WORK...</b>		<b>214</b>
8.1.	INTRODUCTION .....	214
8.2.	THE MAIN FINDINGS RESTATED .....	214

8.3.	IMPLICATIONS .....	216
8.4.	EPISTEMOLOGICAL NOTE .....	220
8.5.	LIMITATIONS OF THE ANALYSIS AND FUTURE WORK .....	222
8.6.	FINAL COMMENTS AND CONCLUSION .....	224
8.6.1	<i>Appendix A: Third Order Utility Expansion</i> .....	226
8.6.2	<i>Appendix B: The Pratt-Arrow Risk Premium</i> .....	229
8.6.3	<i>Appendix C: Portfolio Diversification</i> .....	232
8.6.4	<i>Appendix D: Average Correlation</i> .....	235
8.6.5	<i>Appendix E: A Simple Dynamic Strategy Based on Average Correlation</i> .....	237
<b>BIBLIOGRAPHY .....</b>		<b>238</b>



# Chapter 1: Introduction and Overview

## 1.1. Introduction

Any distribution can be described by its moments. For example, the mean is the first moment, variance is a second moment ‘centered’ around the mean and skewness is a centered third moment. In this thesis, I make large use of the notion of moments in discussing and characterizing the multivariate distribution of equity returns. I study mainly first, second and third moments. I pay special attention to their interaction, as emphasized by modern asset pricing theory, and I discuss the portfolio, investment and risk management implications of alternative first and second moment models, of alternative views about their role in asset pricing and of a number of related empirical findings.

In the next section, I define and contrast unconditional and conditional moments. In Section 1.3, I introduce the old and new paradigms of asset returns. In Section 1.4, I specify the main research questions. In Section 1.5, I explain the motivations of this study. In Section 1.6, I outline the structure of this thesis. In Section 1.7, I provide an overview of the main theoretical results and empirical findings and I highlight their contribution to the extant financial and econometric literature. Section 1.8 concludes.

## 1.2. Conditional vs. Unconditional Moments

Time series of asset returns can be seen as realizations  $y_1, y_2, \dots, y_t$  of a multivariate random variable  $y$  drawn from a joint probability distribution  $p(y_1, y_2, \dots, y_t)$ . Similarly, future returns can be seen as realizations  $y_{t+1}$  of a random variable drawn from the conditional probability distribution  $p(y_{t+1}|y_1, y_2, \dots, y_t)$ . Loosely speaking, stationary series have time invariant moments. Strictly stationary series are realizations of random variables drawn from a time invariant probability distribution and, therefore, all their moments are time invariant. Covariance stationary, also

known as wide sense or weakly stationary, series have finite and time invariant first, second and cross-second moments (e.g., respectively, the mean, variance and autocovariances/autocorrelations). Thus, strictly stationary series with *finite* first and second moments are also covariance stationary but not vice versa. Since independence of two random variables refers to the possibility of writing their joint density function as the product of their marginal densities, serial independence requires that all the cross-moments between any polynomial of current and past realizations be zero. It therefore requires independence between all the moments. Formally, for any random process, and hence also for any return  $y_t$ , serial independence (i.e. independence between  $y_t$  and  $y_{t-i}$ ) means that  $E[g(y_t)h(y_{t-i})] = E[g(y_t)]E[h(y_{t-i})]$  for any integer  $i$ , implying that  $Cov[g(y_t), h(y_{t-i})] = 0$  for any measurable function  $g$  and  $h$  and, therefore, for any cross-moment of  $y_t$  and  $y_{t-i}$ . Autocorrelation is one possible source of serial dependence in returns. It implies linear dependence of the mean of the process on past realizations, and it therefore corresponds to dependence in the first moment. More general forms of serial dependence introduce linear relations between different moments. These might appear as non-linearities in the dependence in first moments.

One way to summarize serial dependence and co-dependence between moments is to let the corresponding polynomials of returns  $g(y_t)$  be determined by data generating processes similar to those commonly used for returns, e.g. auto-regressive moving averages (ARMA). Serial dependence between moments can then be modelled as dependence between polynomials of current returns and past return realizations. For example, with  $g(y_t) = y_t$ , the expectation of  $y_t$  conditional on its past history  $\mu_{y_t} = E_{t-1}(y_t)$ , i.e. the first conditional ‘centered’ moment of  $y_t$ , can be defined as a function of  $y_{t-i}$  ( $i > 0$ ). Using a simple autoregressive specification, we might let  $g(y_t) = a + bg(y_{t-1}) + \varepsilon_t$  or  $y_t = a + bg(y_{t-1}) + \varepsilon_t$ , where  $a$  and  $b$  are constants and  $\varepsilon_t = y_t - \mu_t$  is a conditionally zero-mean return innovation. In this specification, the first moment is a function of the past realization of the process, i.e.  $\mu_{y_t} = E_{t-1}(y_t) = a + by_{t-1}$ . Similarly, with  $g(y_t) = \varepsilon_t^2$ , the conditional expectation



$\sigma_{y_t}^2 = E_{t-1}(\varepsilon_t^2) = (y_t - \mu_t)^2 = (y_t - a - by_{t-1})^2$  is the conditional variance of the return process and it depends upon the past history of the latter.

Specifications like these introduce the distinction between conditional and unconditional moments and allow the former to be time-varying. In Finance, this distinction is important unless we assume that assets are held for a long period of time. In this case the relevant conditioning information set is far in the past and its influence on conditional expectations is negligible, e.g. for (stationary) series for which  $E[g(y_t)]$  exists,  $\lim_{k \rightarrow +\infty} E_{t-k}[g(y_t)] = E[g(y_t)]$ . The distinction between conditional and unconditional variance was emphasized by Engle (1982).

If a process is covariance stationary, its unconditional moments exist and they are the mean of the conditional moments over all the possible realizations of the process itself, i.e.

$$\begin{aligned} E\left\{E_{t-1}[g(y_t)] \mid t-1=1, 2, \dots, +\infty\right\} \\ = E\left\{E_{t-k}[g(y_t)] \mid k \in (-\infty, +\infty)\right\} \\ = E\left\{E_{t-k}[g(y_t)] \mid k=1, 2, \dots, +\infty\right\} \end{aligned} \quad (1.1)$$

Thus, if a process is co-variance stationary, its unconditional variance exists and it is the expectation of the conditional variance conditional upon all the possible realizations, i.e. letting  $g(y_t) = \varepsilon_t^2$ , the unconditional variance of  $y_t$  is

$$\sigma_y^2 = E\left\{E_{t-k}\varepsilon_t^2 \mid k=1, 2, \dots, +\infty\right\} \quad (1.2)$$

As pointed out by Loretan and Phillips (1994), the existence of unconditional moments critically depends on the shape of the density in the tails, i.e. if the density



function does not decline rapidly enough as we move away from the centre of the distribution some of the moments might not exist. For example, returns with finite conditional variance might display infinite unconditional variance. The density in the tails of a distribution, i.e. its ‘tickness’ and the related height of the peak towards the central part of the distribution (leptokurtosis), is captured by the fourth moment, the kurtosis.

### 1.3. Old and New Paradigm

The last decades have witnessed a radical transformation in the way financial theory and financial econometrics researchers model one of their primary objects of interest, asset returns. In the old paradigm, returns were thought to be independently drawn from an underlying joint distribution with time-invariant moments and all the moments existed. In other words, returns were assumed to be independently, identically distributed (henceforth, *i.i.d.*). This view was particularly common in the fifties and sixties and it is summarized by the random walk representation (see Malkiel (1990) for a discussion) of the asset price process with constant drift and white noise error<sup>1</sup>. In the random walk model of asset prices, returns have finite moments of any order and conditional and unconditional moments are the same. Normality of the error term, moreover, implies that the entire multivariate distribution of asset returns can be described by its first and second moments. This, in turn, implies that rational investors should only be concerned about the mean and variance of their portfolios, leaving no room for any role of higher moments in the portfolio optimization problem, as in Markowitz (1952, 1959) mean-variance-portfolio theory. In such a setting, broadly corresponding to the static Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965), only the first two moments of the multivariate return distribution would have asset pricing implications.

---

<sup>1</sup> Formally:  $\frac{dP}{P} = \mu dt + \varepsilon \sigma \sqrt{dt}$  with  $\mu = E\left(\frac{dP}{P}\right)$ ,  $\sigma^2 = E\left(\frac{dP}{P} - \mu\right)^2$ ,  $P$  is the price,  $\varepsilon$  *i.i.d* and  $E(\varepsilon) = 0$ . Here  $E()$  represents both the unconditional and conditional expectation operator.

This paradigm came under intense scrutiny, especially in the 1980s and 1990s. Four main issues about the distribution of asset returns drew the attention of the empirical financial literature, namely whether financial series are independently distributed, whether they are identically distributed over time, whether all the moments of the asset returns distribution exist and whether returns are normally distributed. A large body of evidence, as summarized in Pagan (1996), has since then made clear that, while high frequency returns are virtually serially uncorrelated and lower frequency returns are generally little auto-correlated, there is considerable serial dependence in higher moments. For example, there is overwhelming evidence of conditional heteroskedasticity and time variation in second moments. Furthermore, evidence on return predictability suggests that first moments are time-varying. For example, while monthly returns are generally found to be strikingly unpredictable<sup>2</sup>, there is evidence that annual returns are somewhat predictable and returns at five-year horizons are very predictable (Fama and French (1989) and Cochrane (1999)) using forecasting variables such as the dividend yield, the price earning ratio and other functions of stock prices normalized by an appropriate divisor to make them stationary. This suggests that the mean of the return process is time varying and driven by a slow moving state variable.

Recent studies in the empirical finance literature have reported evidence of two types of asymmetries in the distribution of stock returns. The first is skewness, i.e.  $E_t(\varepsilon_{i,t+1}^3)$ , or asymmetry in the distribution of individual stock returns, which has been reported and studied by numerous authors over the last three decades. See, among others, Simkowitz and Beedles (1978) and Singleton and Wingender (1986). The second type of asymmetry is in the joint distribution of stock returns. One possible source of such an asymmetry is coskewness, i.e.  $E_t(\varepsilon_{i,t+1}\varepsilon_{j,t+1}^2)$ , where  $\varepsilon_{i,t+1}$  and  $\varepsilon_{j,t+1}$  are two zero-mean return innovations. Evidence that stock returns exhibit

---

<sup>2</sup> Monthly and higher frequency stock returns typically have slight, statistically significant predictability with coefficient of determination  $R^2$  of about 1 percent.



some form of asymmetric co-dependence has been reported by several authors in recent years, see for example Erb, Harvey, and Viskanta (1994), Longin and Solnik (2001), Ang and Bekaert (1999, 2002), Ang and Chen (2000, 2002), Campbell, Koedijk, and Kofman (2002), and Bae, Karolyi, and Stulz (2003). The presence of either of these asymmetries violates the assumption of normally distributed portfolio returns, which underlies traditional mean-variance analysis (see Ingersoll (1987)).

Pagan (1996), Campbell, Lo, and MacKinlay (1997) and Cochrane (1999), among many others, provide a summary of the main stylized empirical features of the multivariate distribution of asset returns, such as serial dependence, time variation in first, second and higher moments and non-normality. As these features have become common characteristics of models of asset returns, the old paradigm has been gradually abandoned in favour of a more complex one. In this new paradigm, the multivariate distribution of asset returns cannot be described simply by its first and second moments and conditional and unconditional moments are not in general the same.

#### **1.4. The Fundamental Research Questions**

The research questions that I address in this thesis are both theoretical and empirical in nature. They can be grouped around two closely related themes. The first theme is the asset pricing problem, concerned with the determination of mean returns. The second theme is the dynamic behaviour of second moments.

Concerning the asset pricing problem, an important research question is whether investors are rewarded not only for holding portfolios that perform poorly when aggregate returns are low, as in Sharpe (1964) and Lintner's (1965) CAPM, but also for holding portfolios that perform poorly when volatility is high. One way of reformulating this question is to ask whether asset coskewness, in addition to the covariance with the market portfolio, explains the cross-section of average asset returns. In the empirical investigation of this issue, I focus on the explanatory power

of coskewness for the cross-section of average returns on a particular set of benchmark assets, i.e. the portfolios formed sorting by industry the NYSE, AMEX and NASDAQ stocks included in the database of the Center for Research on Security Prices (CRSP) of the University of Chicago. A further research question related to the asset pricing problem is whether the idiosyncratic component of aggregate volatility plays any role in asset pricing. I empirically investigate this issue using an extensive dataset on Euro area stocks.

Concerning the second moments of equity returns, the main research question is how to model the variation in equity volatility and correlations over time. I study both market and idiosyncratic volatility, and country, industry and firm-level volatility and correlations. In particular, using a largely original Euro area dataset, I investigate long term trends and short run dynamics such as the dependence of these moments on the sign of return innovations. A related research question concerns how the second moment dynamics of large systems of financial variables, such as portfolios with many assets, can best be modelled.

My overall aim is to study the stylized features of the multivariate return distribution relevant in portfolio theory, asset pricing and risk management. Since financial economics is still very far from the full identification of a structural model capable of jointly describing all the moments of the multivariate distribution of assets returns, a reasonable intermediate aim is to estimate reduced form representations that are empirically successful. At a minimum, however, these representations should display the desirable property of not being in conflict with the modern paradigm of the multivariate distribution of asset returns. Since the latter recognizes that returns first and second moments (mainly volatility and correlation) are not static, both conditionally and unconditionally, it is important to study their dynamics over time and to allow for asymmetry in the multivariate return distribution.



## 1.5. Motivations

The investigation of the asset pricing problem is motivated by its profound implications for capital budgeting, portfolio selection and portfolio management. There is ongoing debate on the ability of the CAPM to explain the cross-section of average asset returns. In particular, there is puzzling evidence on the limited ability of theoretically motivated risk factors to drive out the explanatory power of firm characteristics such as size and book-to-market ratio, see Fama and French (1992, 1993, 1995), momentum, as in Jagadeesh and Titman (1993), coskewness (i.e. systematic skewness), as in Harvey and Siddique (2000) and Dittmar (2002), and industry, as in Moskowitz and Grinblatt (1999) and Dittmar (2002). The evidence on the asymmetry of the multivariate distribution of asset returns suggests that, if investors' preferences are not restricted to be defined only over the first two moments, expected returns might depend on higher order odd moments. This possibility motivates the study of the explanatory power of asset coskewness in the cross-section of excess returns, as in Harvey and Siddique (2000), Dittmar (2002) and Post, Levy and Van Vliet (2003, 2005). I focus on the cross-section of portfolios<sup>3</sup> formed sorting stocks according to the industry in which the issuing firm operates because this characteristic has been used less frequently in the extant empirical literature as a sorting criterion and it is known, see Dittmar (2002), for producing a very dispersed (and therefore challenging) cross-section of average returns. This way I minimize a possible 'data snooping' bias (that might arise when fitting the cross-section of more researched and therefore better known portfolios). I use industry portfolios based on CRSP stock data because the latter has become a benchmark dataset for the empirical literature that studies the cross-section of average returns (perhaps because of the high quality of the data, available in a format that facilitates scientific investigation).

The exploration of the role of the idiosyncratic component of aggregate risk in the determination of expected returns is, as suggested by Constantinides (2002), among

---

<sup>3</sup> I thank K. French for making this data publicly available for download.



the most promising research avenues to extend the neo-classical rational asset pricing model based either on a representative agent or on the assumption of complete capital markets. While there are previous contributions that use US data, e.g. Goyal and Santa Clara (2003), there is no previous study that examines this issue in Euro area stock markets. Therefore, in the empirical investigation of the extent to which volatility, both systematic and idiosyncratic, predicts market return I use European stocks data.

The quest for a better understanding of the dynamic behaviour of systematic volatility, aggregate idiosyncratic volatility and equity correlations, the second major theme of this thesis, is motivated by the relevance of these variables for applied portfolio management, risk management and asset pricing. While systematic, market-wide volatility is most important to the holders of well diversified portfolios, both total and idiosyncratic volatility are relevant for incompletely diversified investors. In particular, as remarked by CLMX (2001), the level of aggregate idiosyncratic risk (as measured by average firm and industry-level variance) is important in determining the number of stocks required to achieve a relatively complete diversification whereas the level of aggregate asset correlation determines the extent to which portfolio managers can benefit from diversification. Knowledge of both systematic and idiosyncratic volatility patterns and of the correlation among assets is also relevant to the financial industry since they are important parameters for pricing contingent-claims derivative securities. The relative importance of diversifiable and non-diversifiable risk is also relevant for risk managers and regulators alike, as it is a required input of market and credit risk “internal” models adopted by banks and other regulated intermediaries under the new Basel II Capital Accord<sup>4</sup>. Also, since the level and dynamics of aggregate correlation, both in the short and in the long run, have obvious systemic implications, their knowledge is particularly important for the regulators of financial intermediaries and markets. It is therefore important to investigate whether the developments reported by CLMX (2001), mainly a long-run upward trend in firm-level volatility and declining

---

<sup>4</sup> See on this, for example, the CreditMetrix technical document distributed by JP Morgan.



correlations, are specific to United States markets or whether they obtain more globally. Thus, in studying volatility and correlation dynamics, I will mainly focus on European stocks. This is because, while the seminal contribution of CLMX (2001) has sparked a burgeoning literature focussing on US data, very little is known on the behaviour of idiosyncratic volatility and industry and firm-level correlations trends in the Euro Area.

The study of the behaviour over time of correlations amongst European financial markets is especially important from an investment management perspective. For example, a growing body of empirical evidence on the performance of mutual and pension fund managers suggests that they often under-perform their benchmarks (Blake and Timmerman (1998), Wermers (2000), Baks, Metrick and Wachter (2001), and Coval and Moskowitz, (2001)). Instead of engaging either in expensive but ineffective active portfolio management practices or in almost equally expensive attempts to fully replicate international stock market portfolios, the strategy of buying the market index for each country might yield an effective and cost-efficient international diversification. This provides asset diversity within each country together with international diversification across political frontiers. The success of this strategy, however, depends crucially on the magnitude of the correlations among national markets relative to the correlation among the stocks included in the market index for each country. International correlations tend also to rise with the degree of international equity market integration (Erb, Harvey and Viskanta (1994) and Longin and Solnik (1995)), which has gathered pace in Europe since the mid-1990s (Hardouvelis, Malliaropulos and Priestley (2000) and Fratzschler (2002))<sup>5</sup>. It is therefore of considerable interest to investigate the relative strengths of the trends in correlations at the market index level as well as at the firm level in European equity markets, because the findings have relevance for the diversification properties of passive and active international investment strategies.

---

<sup>5</sup> Fratzschler (2002) also notes that the euro-zone equity market has now surpassed the United States market as the most influential determinant of euro-zone country equity returns.



## 1.6. Structure of the Thesis

The first part of this thesis is devoted to the discussion of the extant literature on asset pricing and volatility. This is done in Chapters 2 and 3, respectively. The second part presents original, largely empirical results on asset pricing and on volatility modelling. In particular, my novel contributions appear in Chapters 4 to 7 and in Chapter 8 I discuss their implications for the asset pricing and investment management problem.

I first discuss, in Chapter 2, the modern view of the multivariate distribution of asset returns within the conceptual framework of modern asset pricing theory. This provides the motivation for the discussion, in Chapter 3, of the theoretical and empirical literature on second moments. Chapter 4 focuses on the cross-sectional dimension of the asset pricing problem and, in particular, on whether coskewness helps explain the cross-section of average returns. This Chapter is based on two papers, Potì (2005b and 2005c), presented at the Annual Meeting of the Financial Management Association (New Orleans, 2004) and discussed at the Doctoral Colloquium of the European Finance Association (Moscow, 2005). In Chapter 5, I study systematic and idiosyncratic volatilities and correlations of Euro area stocks in a simple unconditional setting. This Chapter is based on two studies, Kearney and Potì (2003) and Kearney and Potì (2005b). The former was presented at the European Finance Association (EFA) Annual Meeting (Glasgow, 2003). The latter was presented at the European Financial Management Association (EFMA) Annual Meeting (Milan, 2005) and has been accepted for presentation at the Financial Management Association (FMA) Annual Meeting (Chicago, 2005). In Chapter 6, I model and estimate correlations in the Euro area in a conditional setting. This Chapter is based on a paper forthcoming in the *Research in International Business and Finance*, Kearney and Potì (2005a). Chapter 7 is devoted to the study of the time series of stock market returns and, in particular, of the relation between volatility of discount factors and conditional return volatility. This Chapter draws heavily on a paper forthcoming in the *Applied Financial Economics Letters*, Potì



(2005a). Chapter 8 summarizes the main findings, provides a discussion of their implications, outlines directions for future research and draws together the conclusions. More details on the papers on which this thesis is based, together with sample econometric code (mainly written in RATS, but I will likely add code in other languages in the future) and data sets used for some of the estimation procedures, is available on my website, [www.valeripot.com](http://www.valeripot.com), and on my personal page on the Social Science and Research Network website.

## **1.7. Main Findings and Contributions**

This thesis contributes to the empirical asset pricing literature on both the cross-section and time series of stock returns. It also contributes to the literature on the dependence structure of stock market returns and to the recent but rapidly growing literature on total and idiosyncratic risk. I now provide a detailed account of the main findings and contributions of each chapter.

In Chapter 4, I derive a novel beta-pricing representation of Harvey and Siddique's (2000) version of the 3-moment conditional CAPM. Its main advantage is that both its beta coefficients and risk premia can be interpreted as parameters of appropriate regression models. In an empirical application to US stocks sorted into 30 industry portfolios, I also add to the extant evidence that, while coskewness helps explain a substantial portion of the cross section of average returns and the coskewness premium is of the same order of magnitude as the covariance premium, the estimated unconstrained 3-moment model implies a utility function that is not concave over the relevant wealth range. This confirms earlier results, based on different datasets and a different conditioning information set, reported by Dittmar (2002) and Post, Levy and VanVliet (2003).

In Chapter 5, I extend CLMX's (2001) study to European equity markets. To this end, I construct a unique and comprehensive firm and industry-level dataset of stock returns and volatilities in the Euro area. I also derive in an alternative and more

intuitive manner the main results underlying CLMX's (2001) decomposition of total risk into idiosyncratic and systematic risk and I develop a version of this decomposition based on returns instead of excess returns. Furthermore, I derive an original average correlation index that is computationally easy to construct and I discuss its properties. This average correlation index is very useful to study the aggregate dynamic behaviour of the correlations between a large number of assets. I find that European stocks have indeed become more volatile, and that idiosyncratic risk is the largest component of this volatility. I also find that the potential benefit of diversification strategies in Europe remains substantial and relatively stable over time. Because of the larger idiosyncratic volatility of the typical stock, however, it now takes many more stocks to diversify it away. For example, the number of stocks required for residual idiosyncratic volatility to be reduced to 5 percent in a portfolio of European stocks has risen from 35 in 1974 to 166 in 2003. Stock correlations are on average low, i.e. their long run mean is about 20 percent, thus implying a correspondingly low explanatory power for the market model, i.e. about 4 percent. This is about half the explanatory power of the market model for the average US stock reported by CLMX (2001). CLMX (2001) find, however, that the explanatory power of the market model is declining over time in the United States, while there is no evidence of such a phenomenon in the Euro area. Market volatility forecasts both industry and firm-level volatility. This contrasts with CLMX (2001) who find that firm-level volatility predicts both market and industry-level volatility in the United States. These findings are of interest for investors throughout the world who hold international equity portfolios and especially for European individual and institutional investors who are recently tending to hold greater proportions of their portfolios in stocks<sup>6</sup>.

In Chapter 5, using Euro area data, I also confirm earlier findings based on US data, e.g. Whitelaw (1994) and Goyal and Santa Clara (2003), that the simple lagged

---

<sup>6</sup> The desire to supplement social security benefits and public pension provisions, shrinking because of a rapidly ageing population, contributes towards this shift in investment habits. See Guiso, Haliassos and Jappelli (2002) for an extensive review of the empirical evidence on increasing stock market participation in Europe and the importance of its demographic determinants.



market variance-return relation is positive but statistically insignificant. However, I question Goyal and Santa Clara's (2003) conclusion that the return on the market portfolio is positively related to the idiosyncratic component of lagged total volatility. Instead, I find that when market volatility is also included in a long-run predictive regression, the relation between market return and average idiosyncratic volatility is negative. This is a striking confirmation, obtained using Euro area data, of a similar finding recently reported by Guo and Savickas (2003) for a portfolio of the 500 largest US stocks. I show that this result depends on the circumstance that market and average idiosyncratic volatility jointly proxy for average correlation and thus, perhaps surprisingly, for a component of systematic risk.

In Chapter 6, I estimate the correlations of Euro-area market, industry and sector stock indices in a large-scale multivariate conditional setting, using extensions of Engle's (2002) Dynamic Conditional Correlation GARCH model. For comparison, I also apply the same estimation strategy to a sample of Euro area stock with largest market capitalization. I do not find evidence of a deterministic time trend either in individual stock and industry correlations, consistently with the findings in Chapter 5, or in sector or country level equity correlations. The latter however dramatically increase in recent years due to a structural break shortly before the launch of the Euro. These findings imply that, while industry and firm level diversification have retained their effectiveness, investment managers engaged in asset allocation in the Euro area should not rely any longer on country level diversification strategies. Moreover, the very high correlation between Euro area markets implies that a European-wide factor should largely replace country factors (and perhaps US factors) in factor models of European stock returns. Finally, as reported mainly in Chapter 6, I find weak evidence of increasing correlations depending on past returns, but I confirm the presence of a more general, yet unknown, form of asymmetry in the dependence structure of equity returns in the Euro area, at the firm, industry and country level.

Chapter 7 contains original analytical results on the relation between discount factor volatility and conditional return volatility. This relation provides the amount of discount factor volatility that rational asset pricing models must generate in order to explain conditional return volatility. In an empirical application to the US stock market, I find that it takes about 7 percent discount factor volatility to explain the conditional return volatility from 1871 to 2003, given about 4 percent dividend growth volatility.

## **1.8. Summary and Conclusions**

In this Chapter, I first introduced some preliminary material on the distinction between conditional and unconditional moments and how it arises in the transition from the old to the new paradigm of asset returns. I then specified the fundamental research questions of this thesis and discussed the motivations that drive their investigation. In particular, I discussed the connection between the research questions and unresolved issues in asset pricing and second moment modelling and their relevance for applied portfolio and investment management. Finally, I outlined the structure of the thesis and I summarized the main findings and their contribution to the extant literature.



## **Chapter 2: The Asset Pricing Literature**

### **2.1. Introduction**

In this chapter, I review the literature on the distribution of asset returns and on the closely related topic of asset pricing. Rather than attempting to list all the countless contributions, my aim is to show how the discovery of moment dynamics and their role in asset pricing unfolded over the transition from the old view of asset returns, based on the random walk model and on the identity between conditional and unconditional moments, to the new paradigm that allows for time-varying conditional moments and returns predictability. I first show, however, how all asset pricing models can be derived as specializations of a common analytical framework, the general stochastic discount factor model with possibly time-varying risk premia and returns predictability. I then discuss at more length selected specifications consistent with the modern view of asset returns. In particular, I focus on specifications that allow for systematic skewness and idiosyncratic risk to play a role in asset pricing.

The Chapter is organized in seven main sections. The next two sections introduce the stochastic discount factor approach to asset pricing. In particular, Section 2.2 focuses on the stochastic discount factor framework and its beta-pricing representation and Section 2.3 derives the general linear factor model from the stochastic discount factor representation of the asset pricing problem. In Section 2.4 and 2.5, I provide a brief account of developments in the theory of efficient markets and rational asset pricing, I introduce coskewness as a priced risk exposure, and I discuss the possible role of idiosyncratic risk in asset pricing. Section 2.6 reviews the alternative behavioural approach to asset pricing and Section 2.7 discusses the dichotomy between absolute and relative pricing. The last section summarizes the chapter and draws together the conclusions.

## 2.2. Stochastic Discount Factor Pricing

I begin the discussion of the asset pricing problem imposing at first as little structure as possible. I do this by invoking a theorem credited to Harrison and Kreps (1979). This theorem says that, given free portfolio formation and the law of one price, a stochastic process  $m_{t+1}$  that prices all assets exists. This process is called the stochastic discount factor (henceforth SDF) and satisfies the following condition for all payoffs  $x_{t+1}$  and payoff prices  $p_t$ :

$$p_t = E_t(m_{t+1}x_{t+1}) \tag{2.1}$$

Here, the expectation is taken conditional on the available information set. Under the additional assumption of no arbitrage, as shown by Harrison and Kreps (1979) and popularized by Hansen and Richard (1987),  $m_{t+1}$  must be positive. Moreover, if  $m_{t+1}$  lies in the payoff space, then it is unique. Thus, in complete capital markets, the SDF is unique. We might denote this unique SDF as  $m_{t+1}^*$ . Conversely, if  $m_{t+1}$  is not restricted to lie in the payoff space, any process with the same projection  $m_{t+1}^*$  on the payoff space will satisfy (2.1) and price the payoffs. Thus, in an incomplete capital market, there is an infinite number of  $m_{t+1}$  such that (2.1) holds for every priced payoff  $x_{t+1}$ . In other words, if the set of the assets being priced spans all possible payoffs,  $m_{t+1}$  is unique. Instead, if the priced assets are only a subset of the universe of assets, there is an infinite choice of processes  $m_{t+1}$  that satisfy (2.1). These processes share the same projection on the priced payoff space. Trivially, any linear combination of these processes prices the assets. Virtually all asset pricing models can be derived by defining what determines  $m_{t+1}$ .

### 2.2.1 Alternative Representations

Using the familiar statistical result that the expectation of the product of two random variables equals the product of their expectations plus their covariance, the SDF model can be expressed in a more useful guise as follows:



$$p_t = E_t(m_{t+1}) E_t(x_{t+1}) + Cov_t(m_{t+1}, x_{t+1}) \quad (2.2)$$

By definition, the price of any payoff constructed as the sum of one and the return on an asset, i.e.  $1 + R_{i,t}$ , must be equal to 1. Thus, setting  $p_t = 1$  in (2.2):

$$1 = E_t(m_{t+1}) E_t(1 + R_{i,t+1}) + Cov_t[m_{t+1}, (1 + R_{i,t+1})] \quad (2.3)$$

Therefore,

$$1 + E_t(R_{i,t+1}) = \frac{1}{E_t(m_{t+1})} - \frac{Cov_t(m_{t+1}, R_{i,t+1})}{E_t(m_{t+1})} \quad (2.4)$$

Notice that (2.1) must also apply to the asset with the conditionally risk-free rate of return  $R_{f,t+1}$ . Therefore, setting  $p_t = 1$  and  $x_{t+1} = 1 + R_{f,t+1}$  and noting that, by definition,  $R_{f,t+1}$  is known at time  $t$ , (2.1) implies  $E_t(m_{t+1})(1 + R_{f,t}) = 1$  and therefore:

$$1 + R_{f,t} = \frac{1}{E_t(m_{t+1})} \quad (2.5)$$

Using (2.5), (2.4) can be rewritten in terms of excess-returns<sup>7</sup>, defined as the difference between the asset return and the return on the risk free asset:

---

<sup>7</sup> This relation can also be derived considering that, since excess returns  $r_{i,t}$  are the difference between two unit-price payoffs (the payoff  $1 + R_{i,t+1}$  of asset  $i$  and the payoff  $1 + R_{f,t+1}$  of the risk free asset) their price must be zero under the law of one price. More formally, consider payoffs formed by 1 plus a rate of return, i.e.  $1 + R_{i,t+1}$ . In the terminology of the SDF literature, these payoffs are called ‘gross returns’ whereas  $R_{i,t+1}$  are called ‘net returns’. The price of gross returns is 1 by definition, as you must pay 1 unit of the numeraire (say, a particular currency) to purchase them. Thus, since excess returns can be seen as the difference between two gross returns, i.e. the gross return on a risky asset and the gross return on the risk free asset, their price is zero. Therefore, replacing  $x_{t+1}$  with  $r_{i,t+1}$ , setting  $p_t$  to zero and solving for the expected excess return on asset  $i$  we have the result.

$$E_t(r_{i,t+1}) = \frac{-Cov_t(m_{t+1}, r_{i,t+1})}{E_t(m_{t+1})} \quad (2.6)$$

Using (2.5), the latter can be rewritten as follows:

$$E_t(r_{i,t+1}) = -(1 + R_{f,t+1})Cov_t(m_{t+1}, r_{i,t+1}) \quad (2.7)$$

An interesting implication of (2.7) is that no strategy can offer a (discounted) Sharpe Ratio higher than the volatility of the stochastic discount factor. This is the Hansen and Jagannathan (1991) bound, and it follows from (2.7) because the correlation between any two variables is bounded above and below by one, i.e.

$$Cov_t(m_{t+1}, r_{i,t+1}) = Corr_t(m_{t+1}, r_{i,t+1})\sigma_t(m_{t+1})\sigma_t(r_{i,t+1}) \leq |\sigma_t(m_{t+1})\sigma_t(r_{i,t+1})| \quad \text{and thus}$$

$$\frac{E_t(r_{i,t+1})}{\sigma_t(r_{i,t+1})} \leq (1 + R_{f,t})\sigma_t(m_{t+1}) \quad (2.8)$$

Dividing and multiplying the right hand side of (2.6) by the variance of the SDF yields the beta-pricing representation of the implications of (2.1) for excess returns:

$$E_t(r_{i,t+1}) = \beta_{m,i,t} \left[ -\frac{Var_t(m_{t+1})}{E_t(m_{t+1})} \right] \quad (2.9)$$

Here,  $\beta_{m,i,t} = \frac{Cov_t(m_{t+1}, r_{i,t+1})}{Var_t(m_{t+1})}$  is a coefficient of the regression of the asset excess-return on the SDF. Using (2.5), the beta-pricing representation in (2.9) can be rewritten as follows:

$$E_t(r_{i,t+1}) = \beta_{m,i,t} [-Var_t(m_{t+1})(1 + R_{f,t})] \quad (2.10)$$



### 2.2.2 Excess Returns and the Mean of the SDF

For realistically low values<sup>8</sup> of  $R_{f,t+1}$ , we can approximate the right-hand side of (2.5) as follows:

$$\frac{1}{E_t(m_{t+1})} = 1 + R_{f,t} \cong 1 \quad (2.11)$$

Thus, for realistically low values of the risk free rate, the mean of the SDF is approximately equal to 1. Under this approximation, the mean of the SDF does not identify excess return means in (2.7) and (2.10). Using (2.11) in (2.7), then we might approximate expected excess returns as follows:

$$E_t(r_{i,t+1}) \cong -Cov_t(m_{t+1}, r_{i,t+1}) \quad (2.12)$$

Similarly, using (2.11) in (2.10), we might approximate the beta pricing representation of expected excess returns in the following way:

$$E_t(r_{i,t+1}) \cong \beta_{m,t} [-Var_t(m_{t+1})] \quad (2.13)$$

### 2.3. Factor Models

Equation (2.1) and its representations in (2.7) and (2.10), like their approximate counterparts in (2.12) and (2.13), represent very general asset pricing results. All conditional linear factors models can be derived as specializations of these equations by specifying  $m_{t+1}$  as a linear function of a number of factors  $f_{t+1}$ :

$$m_{t+1} = a_t + b'_t f_{t+1} \quad (2.14)$$

---

<sup>8</sup> The historical average of the risk-free rate is about 6 percent per annum, see for example Cochrane (2001).

This is said to be a conditional linear factor model (LFM) because the parameters of the SDF are allowed to be time-varying conditional on available information. Fixing the parameters of the SDF yields instead the general unconditional LFM. The asset pricing implications of any conditional and unconditional linear factor model can be represented as specifications of the implications of (2.1) and (2.14). Using the latter in (2.4), the no-arbitrage implications for the cross-section of expected returns are:

$$1 + E_t(R_{i,t+1}) = \frac{1}{E_t(m_{t+1})} - \frac{Cov_t(R_{i,t+1}, f_{t+1})b_t}{E_t(m_{t+1})} \quad (2.15a)$$

Or,

$$\begin{aligned} E_t(R_{i,t+1}) &= R_{f,t+1} - (1 + R_{f,t+1})Cov_t(R_{i,t+1}, f_{t+1})b_t \\ &= R_{f,t+1} + (1 + R_{f,t+1})Cov_t(R_{i,t+1}, f_{t+1})\xi_t \end{aligned} \quad (2.15b)$$

Here,  $Cov_t(R_{i,t+1}, f_{t+1})$  can be interpreted as the quantity of factor risk and  $\xi_t = -b_t$  as the vector of factor risk prices. This equation can be easily rewritten in beta-pricing form, thus giving the beta-pricing representation of the asset pricing implication of (2.1) and (2.14):

$$E_t(R_{i,t+1}) = R_{f,t+1} + \beta'_{i,t}\lambda_t \quad (2.16)$$

where,

$$\beta_{i,t} = Var_t(f_{t+1})^{-1}Cov_t(f_{t+1}, R_{i,t+1}) \quad (2.17)$$

$$\lambda_t = -(1 + R_{f,t+1})Var_t(f_{t+1})b_t \quad (2.18)$$



Here,  $\beta_{i,t}$  is a vector of coefficients from the regression of asset  $i$  on the factors and  $\lambda_t$  is a parameter vector. The former can be seen as the factor loadings and the latter can be seen as the risk premia or the price of the factors. When pricing excess returns, we can use the approximation in (2.11) and set  $\frac{1}{E_t(m_{t+1})} = a_t + b_t' E_t(f_{t+1}) \cong 1$ . Then we can apply (2.12) and (2.13) and rewrite the

covariance and beta-pricing relations in (2.15b) and (2.16) as follows:

$$E_t(r_{i,t+1}) = -Cov_t(r_{i,t+1}, f_{t+1})(1 + R_{f,t})b_t \cong -Cov_t(r_{i,t+1}, f_{t+1})b_t \quad (2.19)$$

$$E_t(r_{i,t+1}) = \beta_{i,t}' \lambda_t \quad (2.20)$$

where,

$$\beta_{i,t} = Var_t(f_{t+1})^{-1} Cov_t(f_{t+1}, r_{i,t+1}) \quad (2.21)$$

$$\lambda_t = -(1 + R_{f,t})Var_t(f_{t+1})b_t \cong -Var_t(f_{t+1})b_t \quad (2.22)$$

Here,  $\beta_{i,t}$  is a vector of coefficients from the regression of asset  $i$  excess returns on the factors while all the other symbols are defined as before.

Letting  $z_t$  represent a vector of variables that summarize the relevant conditioning information, we can write  $a_t = a_t(z_t)$  and  $b_t = b_t(z_t)$  for conditional models. The simplest way to model conditional time-variation in the parameters of (2.14) is to specify them as a linear function of the set of conditioning variable:

$$\begin{aligned} a_t &= a^0 + a^1 z_t \\ b_t &= b^0 + b^1 z_t \end{aligned} \quad (2.23)$$

Here,  $a_t$  is a scalar,  $b_t$ ,  $b^0$  are  $N \times 1$  vectors,  $z_t$  is a  $k \times 1$  vector of conditioning variables,  $a^1$  is a  $k \times 1$  vector and  $b^1$  is an  $k \times N$  matrix. Using, (2.23) we can rewrite (2.14) as follows:

$$\begin{aligned}
 m_t &= a^0 + a^1 z_t + (b^0 + b^1 z_t)' f_{t+1} \\
 &= a^0 + a^1 z_t + b^0 f_{t+1} + z_t' b^1 f_{t+1} \\
 &= a^0 + a^1 z_t + b^0 f_{t+1} + b^2 (f_{t+1} \otimes z_t)
 \end{aligned} \tag{2.24}$$

Here,  $b^2$  is a  $[(k \times N) \times 1]$  vector obtained stacking the  $N$  columns of  $b^1$ , i.e

$$b^2 = \text{vec}(b^1) \tag{2.25}$$

The specification in (2.24) and (2.25) can be seen as an unconditional model, i.e. a model with time-invariant parameters, in the new set of factors

$$F_t = \begin{bmatrix} z_t \\ f_{t+1} \\ f_{t+1} \otimes z_t \end{bmatrix} \tag{2.26}$$

For convenience, we can also rewrite (2.24) folding the unconditional mean of these factors in the constant and write the SDF as a linear function of a new set of unconditionally de-measured factors:



$$\begin{aligned}
m_t &= [a^0 + a^1 E(z_t) + b^0 E(f_{t+1}) + b^2 E(f_{t+1} \otimes z_t)] + \\
&\quad + a^1 [z_t - E(z_t)] + b^0 [f_{t+1} - E(f_{t+1})] \\
&\quad + b^2 [f_{t+1} \otimes z_t - E(f_{t+1} \otimes z_t)] \\
&= \tilde{a} + a^1 \tilde{z}_t + b^0 \tilde{f}_{t+1} + b^2 (f_{t+1} \tilde{\otimes} z_t) \\
&= \tilde{a} + b^3 \tilde{F}_t
\end{aligned} \tag{2.27}$$

Here,

$$\tilde{a} = a^0 + a^1 E(z_t) + b^0 E(f_{t+1}) + b^2 E(f_{t+1} \otimes z_t) \tag{2.28}$$

$$\tilde{F}_t = \begin{bmatrix} z_t - E(z_t) \\ f_{t+1} - E(f_{t+1}) \\ f_{t+1} \otimes z_t - E(f_{t+1} \otimes z_t) \end{bmatrix} \tag{2.29}$$

$$b^3 = \begin{bmatrix} a^1 \\ b^0 \\ b^2 \end{bmatrix} \tag{2.30}$$

Since the parameters of the factors in (2.27) are by definition constant over time, the conditional and unconditional implications of the model are the same. In particular, we can derive the unconditional implications without worrying about co-variation between the parameters of the SDF and the factors. To this end, we may take the unconditional expectation of (2.1) with the SDF specified as in (2.27). The covariance and beta-pricing representations of the implication of this unconditional expectation are the following:

$$E(r_{i,t+1}) = -Cov(r_{i,t+1}, \tilde{F}_{t+1}) (1 + r_{f,t}) b^3 \cong -Cov(r_{i,t+1}, F_{t+1}) b^3 \tag{2.31}$$

$$E(r_{i,t+1}) = \beta_i' \lambda \tag{2.32}$$

where,

$$\beta_i = \text{Var}(\tilde{F}_{t+1})^{-1} \text{Cov}(\tilde{F}_{t+1}, r_{i,t+1}) = \text{Var}(F_{t+1})^{-1} \text{Cov}(F_{t+1}, r_{i,t+1}) \quad (2.33)$$

$$\lambda = -(1 + R_{f,t}) \text{Var}(\tilde{F}_{t+1}) b^3 \cong -\text{Var}(F_{t+1}) b^3 \quad (2.34)$$

## 2.4. Asset Pricing Paradigms and EMH

The efficient market hypothesis (henceforth, EMH), as formulated by Fama (1970, 1976), requires that conditional future cash-flows expectations and conditional moments of the multivariate return distribution be formed using all the available *relevant* information<sup>9</sup> and, ultimately, that they do not deviate in any systematic (i.e. exploitable) way from the unconditional ones<sup>10</sup>. This implies that returns deviations from their possibly time-varying equilibrium conditional expectations follow a *fair game* process (see, for a simple taxonomy, Copeland and Weston (1991)) with zero conditional and unconditional mean but with possibly time-varying higher order moments. On average, then, returns equal conditional expected returns or, equivalently, expected returns conditional on the available information set are unbiased estimates<sup>11</sup> of actual future returns. The key difference between rational asset pricing under the old and new paradigm<sup>12</sup> is that conditional expected returns and higher order moments of the returns deviations from their conditional means are fixed in the former and possibly time-varying in the latter.

<sup>9</sup> This is a definition of market efficiency implied by Fama's (1970) discussion and reported in Fama (1976).

<sup>10</sup> Recall that, if the distribution of the relevant conditioning variables is known, unconditional moments can be derived from the conditional ones. Therefore, if either asset cash-flows conditional expectations or conditional return moments are not formed using all the *available* relevant information, superior forecasts of asset prices could be formed by using conditioning variables that convey the relevant information neglected by market prices. These forecasts would be exploitable to earn above-average risk-adjusted returns. Clearly, this does not need to apply to conditional asset cash-flows expectations and return moments formed using subsets of the available information set, such as the data available to the econometrician.

<sup>11</sup> This condition can be formulated as follows:  $E(r_{i,t}) = E[r_{i,t} - E(r_{i,t} | \Omega)]$ , where  $\Omega$  denotes the conditioning information set.

<sup>12</sup> The key difference is therefore that the former relies on a random walk whereas the latter is based on a fair game view of conditionally unexpected returns. Also recall that conditional and unconditional moments are the same only in the random walk case.



The old paradigm implied the EMH, but the reverse is not true. In particular, within the new paradigm of asset returns, it is possible to recognize explanations for asset pricing phenomena based either on asset pricing models with investors that process information and decide upon it rationally, and thus consistently with the EMH, or on models that allow for some degree of investors' irrationality. I call the former rational asset pricing models and the latter behavioural asset pricing models.

## 2.5. Rational Asset Pricing Models

Rational asset pricing models can be interpreted as specifications of a unified theoretical framework, the neoclassical rational economic model (Constantinides (2002)), that views expected excess returns as the reward demanded by risk averse, expected utility optimizing investors for bearing non diversifiable risk. These investors have unambiguously defined preferences over consumption. If we add the assumptions that investors' expectations are rational and investors' beliefs consistent, in the sense implied by Sargent's (1993) discussion of the rational expectation equilibrium, this framework implies the EMH. Versions of this theory allow for market incompleteness, market imperfections, informational asymmetries, and learning. The theory also allows for differences among assets for liquidity, transaction costs, tax status, and other institutional factors.

### 2.5.1 Consumption Asset Pricing

For inter-temporal utility maximizing investors,  $m_{t+1}^j$  depends on their impatience and on the marginal utility of whatever they must give up in order to acquire additional units of the payoff  $x_{t+1}$ . To see this, suppose that investors extract utility from consumption, and that they have the following 2-period utility function:

$$u(c_t^j, c_{t+1}^j) = u(c_t^j) + \beta E[u(c_{t+1}^j)] \quad (2.35)$$

Here,  $c_t^j$  denotes investor  $j$ 's consumption and  $\beta$  is the subjective discount factor that represents the investors' impatience<sup>13</sup>, that is by how much, under *any* circumstance, *any* payoff is worth less if it is paid at a later date. Subjective discount factors should be always less than unity for impatient investors. Desirable properties of investors' utility function, as argued by Arrow (1971), are positive and decreasing marginal utility of wealth and non-increasing absolute risk aversion. Positive marginal utility of wealth, or  $u' > 0$ , implies investors' non satiation (NS), whereas decreasing marginal utility,  $u'' < 0$ , implies risk aversion (RA). Non increasing absolute risk aversion (NIARA),  $\frac{d(-u''/u')}{dc^j} \leq 0$ , implies that risky assets are not inferior goods, and as shown in Arditti (1967), it is a sufficient condition for  $u''' \geq 0$ . Hence  $u''' \geq 0$  implies NIARA and aversion to negative skewness. NIARA, for a utility maximizing, risk-averse individual, and hence with positive marginal utility and RA, is also related to prudence as defined by Kimball (1990). Included in the set of utility functions that display these desirable attributes are the logarithmic, power and negative exponential utility function. It should be noted that the popular quadratic utility function does not satisfy NIARA.

The investor in  $t$  must decide how much to consume and how much to invest in the asset that offers the pay-off  $x_{t+1}$ . Subject to his inter-temporal budget constraint, the more of the asset he purchases, the less his consumption today but the more he will be able to consume in the future. The problem of a rational investor, therefore, is to find the level of investment that maximizes his expected utility. Assuming that the utility function is concave, denoting by  $u'(c_t^j)$  the marginal utility of consumption

and setting  $m_{t+1}^j = \beta \frac{u'(c_{t+1}^j)}{u'(c_t^j)}$ , (2.1) can be seen as the first order condition for the

maximization of the investor's expected utility, i.e. the expectation of (2.35), given the price  $p_t$  of the pay-off  $x_{t+1}$ . In this setup, treating the subjective discount factor as

---

<sup>13</sup> It has nothing to do with the CAPM asset beta but I keep this notation because it is almost standard in the literature.



an inter-temporal constant, the SDF  $m_{t+1}^j$  is proportional to marginal utility growth and (2.7) implies that investors are willing to pay more for assets that are expected to pay off handsomely when marginal utility of consumption is high.

The SDF, i.e. the process  $m_{t+1}$  that prices all pay-offs, depends in general on the circumstances (factors) that determine the extent to which investors' aggregate marginal utility in  $t+1$  is high relative to the previous period. The shape of investors' utility and the extent to which investors can freely form portfolios has also implications for the shape of the SDF that prices the assets. For example, NS imply no-arbitrage and therefore a positive SDF. Furthermore, if the utility function is concave, marginal utility is high when resources to purchase additional units of consumption are scarce and therefore consumption is low. A payoff that would make additional resources available when these are needed the most would be particularly welcome and the investors would value it more ( $m_{t+1}$  is high). This implies a SDF decreasing in wealth. At a more technical level, the shape of investors' utility also has implications for how closely the SDF that prices all assets resembles the shape of individual investors' marginal utility growth. In other words, whether aggregation of individual investors' marginal utility growth results in a SDF defined over aggregate wealth with the same shape as the individual investors' SDF depends, in general, on the shape of utility. In empirical applications, the assumption that prices are set by a representative investor allows to bypass this issue (essentially, leaving it in the background for asset pricing theorists).

### 2.5.2 *Representative Investor*

Under the representative investor assumption,  $c_t^j = c_t$  and the SDF  $m_{t+1}$  can be expressed in terms of marginal utility of aggregate consumption:

$$m_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)} \tag{2.36}$$

The asset pricing implications of the representative investor assumption and of the assumption that capital markets are complete are the same. This is because, in complete capital markets, as in Lucas (1978), investors can exchange contingent claims on any future state of the world. Full risk sharing and diversification are therefore optimal for all investors, who then hold portfolios with risky assets in identical proportions. In these circumstances, pricing assets with respect to individual investors' consumption or with respect to aggregate consumption is equivalent because marginal utility growth is the same for all investors.

In a 2-period setting, investors must consume at the end of the second period all their wealth. Thus, in the SDF in (2.36), we can substitute out the representative investor's consumption with wealth. In a multi-period setting, consumption and wealth are equivalent only if either returns are unpredictable, as in the old paradigm, or predictability has no effect on inter-temporal optimal consumption-investment and portfolio choices. Strictly, the latter condition requires the assumption of logarithmic<sup>14</sup> utility. The empirical literature, e.g. Jagannathan and Wang (1996) and Lettau and Ludvigson (2000), however, often assumes that the SDF pricing equation holds conditionally period by period even under other type of utility functions. This corresponds to the assumption that predictability is at most a second order effect relative to the variability in consumption and wealth. Under these conditions, the inter-temporal marginal rate of substitution in (2.36) can be expressed as a function of aggregate wealth:

$$m_{t+1} = \beta \frac{u'(W_{t+1})}{u'(W_t)} \quad (2.37)$$

The SDF of a representative investor with preferences defined over wealth that display NS, RA and DIARA is positive, decreasing and concave in wealth.

---

<sup>14</sup> As explained in Cochrane (2005), for this type of utility function substitution effects (higher expected return implies an higher opportunity cost of current consumption and therefore tends to decrease it) and income effects (higher expected return imply higher next period wealth and therefore tends to increase consumption) exactly offset each other.



### 2.5.3 CAPM

The CAPM is a special single factor model. In its original version, it is a static equilibrium model. Under investors' NS, it can be derived either assuming a representative investor with quadratic utility, thus excluding preference for moments of the multivariate distribution of asset returns higher than the second, or allowing for preference for higher moments (as under a power utility function) but assuming that returns are multivariate normal, and that investors, rational and risk averse, can freely diversify and have access to the same information. The latter assumption, even when a subset of investors is imperfectly rational, can be replaced by the assumption that the most informed marginal investor is rational and can borrow and lend without limits at the risk-free rate (this, essentially, requires a frictionless capital market). Quadratic utility assures that the  $u'''$  term in (A.5) in Appendix A is zero. Under a multivariate normal distribution, the covariance with the squared rate of return on investor's wealth is zero (because of the symmetry of the normal distribution). In either case, the SDF depends linearly only on the return on the mean-variance efficient portfolio of risky assets, i.e. (2.14) becomes:

$$m_{t+1} = a + br_{m,t+1} \quad (2.38)$$

From (2.32), then, the expected excess return on any asset is proportional to the coefficient of the regression of the asset excess returns on the portfolio excess return and captures the asset systematic risk exposure. The proportionality coefficient, i.e. the risk-premium, is the market expected excess return. This is because, by construction, the regression coefficient  $\beta_{m,m}$  of the market excess return on itself equals 1 and therefore, from (2.32),  $\lambda_m = E(r_{m,t+1})$ .

#### 2.5.4 (C)CAPM

We might extend the CAPM to an inter-temporal setting, where returns are not i.i.d. and moments are allowed to be time-varying, by letting the CAPM hold conditionally, period by period. This is clearly an approximation, as a rational mean-variance investor would anticipate the possibility of variation in the first moment of the return distribution and thus would seek to hedge against adverse (negative) changes in expected returns, i.e. a demand for hedging against reinvestment risk would arise and a corresponding risk premium would enter the equilibrium expected return determination equation. Since in a two-period CAPM expected returns are proportional to expected market variance, the latter would show up as an additional risk factor with a positive risk price in the SDF of the representative investor, as in Merton's (1973) Inter-temporal CAPM (henceforth, ICAPM), i.e.

$$E_t(r_{i,t+1}) = -b_{1,t} Cov_t(r_{m,t+1}, r_{i,t+1}) - b_{2,t} Cov_t(z_{t+1}, r_{i,t+1}) \quad (2.39)$$

Here  $z_{t+1}$  is a state variable that describes the state of the investment opportunity set, i.e. it captures reinvestment risk. Merton (1980), however, points out that the hedging motive is likely not very important. Following Merton's (1980) advice, Jagannathan and Wang (1996) set the price of reinvestment risk to zero and approximate the SDF as a linear function of the return on the market portfolio with time-varying parameters, i.e.  $m_{t+1} = a_t + b_t r_{m,t+1}$ . Such a SDF summarizes the asset pricing implication of the conditional CAPM, henceforth (C)CAPM. Alternatively, we could just treat this specification of the SDF as a reduced form representation of the true inter-temporal SDF. In any case, letting the SDF parameters depend linearly on the conditioning variable  $z_t$ , as in (2.23), and setting  $f_{t+1} = r_{m,t+1}$  in (2.14), we have:

$$\begin{aligned} m_{t+1} &= a_0 + a_1 z_t + (b_0 + b_1 z_t) r_{m,t+1} \\ &= a_0 + a_1 z_t + b_0 r_{m,t+1} + b_1 z_t r_{m,t+1} \end{aligned} \quad (2.40)$$



Using (2.40) in (2.20), the beta-pricing representation of the conditional excess return pricing implications of the (C)CAPM is the following:

$$E_t(r_{i,t+1}) = \beta_{m,i,t} \lambda_{m,t} \quad (2.41)$$

Here,  $\beta_{m,i,t}$  is a time-varying coefficient of the regression of  $r_{i,t}$  on  $r_{m,t}$  and  $\lambda_{m,t}$  is the conditional market risk premium, given by (2.22):

$$\lambda_{m,t} = -Var_t(r_{m,t+1})b_t \quad (2.42)$$

Also, since by definition  $\beta_{m,m,t} = 1$ , we have from (2.41) that  $\lambda_{m,t} = E_t(r_{m,t+1})$ . Hence, the conditional market risk premium is equal to the conditional market expected excess return. Since, as shown in Appendix A,  $b_t = -RRA_t$ , the market risk premium in (2.42) can be rewritten as follows:

$$\lambda_{m,t} = E_t(r_{m,t+1}) = RRA_t Var_t(r_{m,t+1}) \quad (2.43)$$

$RRA_t$  can be interpreted as the representative investor's relative risk aversion parameter for reasons that become clear by examining the derivation of the stylized risk-return relation reported in B.5.

To derive the unconditional implications of the conditional SDF model in (2.40), I apply (2.41) and take unconditional expectations of both sides:

$$E(r_{i,t+1}) = E(\beta_{m,i,t} \lambda_{m,t}) = \beta_{i,t} \lambda_t = \beta_{z,i} \lambda_z + \beta_{m,i} \lambda_m + \beta_{z_m,i} \lambda_{z_m} \quad (2.44)$$

Here,  $\beta_{z,i}$ ,  $\beta_{m,i}$  and  $\beta_{z_m,i}$  are regression coefficients of  $r_{i,t+1}$  on, respectively,  $z_t$ ,  $r_{m,t+1}$  and  $z_t r_{m,t+1}$ . Equivalently, the SDF in (2.40) can be seen as a linear function of  $z_t$ ,  $r_{m,t+1}$  and  $z_t r_{m,t+1}$ . Hence, (2.44) can be derived applying (2.32) to (2.40) with the

elements of  $F_t$  in (2.26) given by  $z_t$ ,  $r_{m,t+1}$  and  $z_t r_{m,t+1}$ . Notice that, if the parameters of the SDF are fixed,  $a_t = a_0 = a$  and  $b_t = b_0 = b$ , the preceding equations simplify to the unconditional CAPM (CAPM), i.e.  $m_{t+1} = a + br_{m,t+1}$ , and  $E(r_{i,t+1}) = \beta_{m,i} \lambda_m$ .

### 2.5.5 Conditioning Variables

A critical consideration in estimating the (C)CAPM, or any conditional asset pricing model, is the choice of the conditioning variable  $z_t$ . The conditioning variable should capture the time variation in the parameters of the SDF. There are two main theoretical reasons why the parameters of a SDF conditionally defined over market wealth might change over time.

One relates to non-market sources of risk and the impact of economy-wide shocks on the marginal utility of stock market wealth. From this perspective, we seek conditioning variables that proxy for the state of the economy and, in particular, for sources of systematic variation in non-market wealth, such as labor income shocks and real estate returns. These are labeled by Cochrane (2001) ‘distress risk’ factors or recession variables and should capture sources of systematic risk different from the stock market. During a recession unemployment is high, labor income is low and more volatile and property prices falter. If investors’ marginal utility of stock market wealth is higher under these circumstances than in good times, variables that capture the state of the economy should show up as priced risk factors alongside the stock market factor. This ultimately implies that investors’ utility is not defined only over stock market wealth but also over other forms of wealth. In turn, this implies that the stock market is not a good proxy for overall wealth. The recession state variables do not need to predict anything (either the stock market or the future state of the economy) but they should be highly correlated with the wider economy or particular (sizeable) portions of it unrelated to the stock market. In other words, they should represent good instruments for the state of portions of the economy unrelated to the stock market but relevant in determining investors’ marginal utility. High correlation implies that the conditioning variable should be either highly pro-cyclical or anti-



cyclical relative to these portions of the economy. If they were pro-cyclical, they would command a positive risk premium. If they were anti-cyclical, they would command a negative risk premium (exposure to them would represent an insurance against a non-stock market source of systematic risk).

The other theoretical reason why the parameters of the SDF might change over time relates, in Merton's (1973) ICAPM framework, to inter-temporal risk and to the impact of changes to the future investment opportunity set on marginal utility of wealth. Thus we seek conditioning variables that summarize the predictable evolution of the investment opportunity set and hence provide a summary measure of expected excess returns. These variables should, in other words, predict excess returns. In particular, in a world where only systematic risk matters to investors, the conditioning variable should help forecast market returns. The empirical literature has proposed a number of variables that help predict future returns. The most successful are the stochastically de-trended short term interest rate, employed among others by Scruggs (1998), the book to market value ratio, the dividend-price ratio, used by Campbell and Shiller (1988), and the observable proxy for the consumption-wealth ratio proposed by Lettau and Ludvigson (2001). Theoretical arguments that suggest that the consumption-wealth ratio and the dividend-price ratio should predict future returns are especially compelling.

To show that the consumption–aggregate wealth ratio summarizes agents' expectations of future returns, Lettau and Ludvigson (2001), using a log-linear approximation to a representative investor's inter-temporal budget constraint  $W_{t+1} = (1 + R_{m,t+1})(W_t - C_t)$ , express the log consumption–wealth ratio in terms of future returns to the market portfolio and future consumption growth. Because this approximation is based on the agent's inter-temporal budget constraint, it holds both ex post and ex ante. Accordingly, the log consumption–wealth ratio may be expressed in terms of expected returns to the market portfolio and expected consumption growth as:

$$c_t - w_t \cong E_t \sum_{j=1}^{\infty} \rho_w^j (r_{m,t+j} - \Delta c_{t+j}) \quad (2.45)$$

Here, lower case letters denote logarithms of (per capita) consumption and wealth and  $\rho_w$  is the steady-state ratio of invested to total wealth. This essentially means that, given the representative investor's wealth, the amount of consumption today depends on the amount he wishes to be able to afford to consume tomorrow and, therefore, on his expected future consumption. Under Muth's (1961) rational expectations (henceforth RE), the above equation implies that, if consumption growth is not too volatile (something that appears to be true empirically), the variation in the log consumption–wealth ratio must be driven by variation in expected returns. It therefore summarizes expectations of future returns on the market portfolio. Intuitively, if the consumption-wealth ratio is high, then the agent must be expecting either high returns on wealth in the future or low consumption growth rates (boosting in both cases current consumption). Since consumption growth rates are fairly stable, however, swings in the consumption-wealth ratio should be related to changing agents' expectations about aggregate returns and, under RE, they should predict aggregate returns.

Of course, the log consumption–aggregate wealth ratio is not observable because human capital is not observable. To overcome this obstacle, Lettau and Ludvigson (2001) construct a proxy based on *observable* quantities. Denote non-human or asset wealth by  $A_t$  and its log as  $a_t$ . Also, assume that human capital  $H_t$  is on average a constant multiple of labor  $Y_t$  income. Its logarithm then can be written as  $h_t = k + y_t + v_t$ , where  $k$  is a constant and  $v_t$  is a mean zero stationary random variable. Lettau and Ludvigson (2001) reformulate the bivariate cointegrating relation between  $c_t$  and  $w_t$  in the consumption-wealth ratio equation ( $c_t$  and  $w_t$  are both integrated but their linear combination on the right hand side is stationary) as a trivariate co-integrating relation involving the three *observable* variables log consumption  $c_t$ , log nonhuman or asset wealth  $a_t$ , and log labor earnings  $y_t$ . Since  $c_t$  and  $a_t$  are both I(1), such a reformulation is possible, by Engle and Granger (1987)



representation theorem, under the condition that labor income is integrated and the rate of return to human capital is stationary. Aggregate wealth is  $W_t \equiv A_t + H_t$  and log aggregate wealth may be approximated as  $w_t \cong \omega a_t + (1 - \omega)h_t$  where  $\omega$  equals the average share of nonhuman wealth in total wealth,  $\frac{A_t}{W_t}$ . The left-hand side of (2.45) may then be expressed as follows:

$$\begin{aligned}
 c_t - w_t &\cong c_t - \omega a_t - (1 - \omega)h_t \\
 &= c_t - \omega a_t - (1 - \omega)(k + y + v)_t \\
 &= c_t - \omega a_t - (1 - \omega)y_t - (1 - \omega)(k + v)_t \\
 &= cay_t - (1 - \omega)k_t - (1 - \omega)v_t
 \end{aligned} \tag{2.46}$$

Here,  $cay_t \equiv c_t - \omega a_t - (1 - \omega)y_t$  is the difference between log consumption and a weighted average of log asset wealth and log labor income. Solving (2.46) for  $cay_t$  and using (2.45), we can write:

$$\begin{aligned}
 cay_t &\equiv (1 - \omega)k + E_t \sum_{j=1}^{\infty} \rho_w^j (r_{m,t+j} - \Delta c_{t+j}) + (1 - \omega)v_t \\
 &\equiv const. + E_t \sum_{j=1}^{\infty} \rho_w^j (r_{m,t+j} - \Delta c_{t+j}) + (1 - \omega)v_t
 \end{aligned} \tag{2.47}$$

Because all the variables on the right-hand side of the above equation are stationary, the model implies that  $cay_t$  is stationary and hence that consumption, asset wealth, and labor income share a common stochastic trend (they are cointegrated), with  $\omega$  and  $1 - \omega$  parameters of this shared trend. If the cointegrating parameter  $\omega$  can be consistently estimated,  $cay_t$  can be treated as observable. As long as the error term  $v_t$  on the right-hand side is not too variable, this equation also implies that  $cay_t$  should be a good proxy for the unobservable quantities on the right hand side of (2.47) and therefore for variation in the log consumption–aggregate wealth ratio and expected returns. An important issue in using the left-hand side of this equation as a

conditioning variable is the estimation of the parameters in  $cay_t$ . Lettau and Ludvigson (2001) discuss how the cointegrating parameter  $\omega$  can be estimated consistently. As suggested by Lettau and Ludvigson (2001), it is the  $cay_t$  time-series constructed using the estimated  $\omega$  parameter and the observed log consumption  $c_t$ , log asset wealth  $a_t$  and log labor earnings  $y_t$  that can be employed as a scaling variable in a conditional asset pricing model.

The specification of the consumption-wealth ratio equation reported above is analogous to the linearized formula for the log dividend–price ratio (Campbell and Shiller (1988)), where consumption enters in place of dividends and wealth enters in place of price:

$$p_t - d_t \equiv const. + E_t \sum_{j=1}^{\infty} \rho^{j-1} (\Delta d_{t+j} - r_{m,t+j}) \quad (2.48)$$

Here,  $d_t$  denotes log-dividends,  $r_t$  denotes returns, and  $\rho$  can be seen as the steady state dividend yield. Because all the variables on the right-hand side of the above equation are stationary, the model implies that  $p_t - d_t$  is stationary and hence that prices and dividends share a common stochastic trend (they are cointegrated), with 1 and  $-1$  parameters of their cointegrating relation. If the dividend-price ratio is high, investors must be expecting either high returns on the stock market portfolio in the future or low dividend growth rates. Since both consumption and dividends are not very volatile and their growth rates are relatively unpredictable, high wealth and high stock market prices relative to, respectively, consumption and dividends (but also relative to the book value and other metrics) must predict low future returns. The key difference between the consumption-wealth ratio and the dividend-price ratio is what is on the right-hand side: in the equation for the consumption-wealth ratio it is the return to the *entire* market portfolio and consumption growth, whereas in the dividend-price ratio equation it is the return to the stock market component of wealth and dividend growth.



Lettau and Ludvigson (2001) and Guo and Savickas (2003) present evidence that  $cay_t$  is a good predictor of excess returns on aggregate stock market indices. Evidence that the price-dividend ratio is a good predictor of returns is given, among others, by Campbell and Shiller (1988), Campbell (1991) and, more recently, Cochrane (1999, 2001). It is worth stressing that predictability is a long-horizon effect. The predictability of 1 to 5 year returns using the dividend-price ratio as a forecasting regression variable is reported in Table 2.1, reproduced from Cochrane (1999). The dividend-price ratio predicts 17% of the variation in 1 year returns and its explanatory power rises steadily as the horizon increases. It predicts up to 59% of the variation in 5 year returns. Even though the explanatory power,  $R^2$ , of the regression is inflated by an overlapping observations problem, the results at different horizons are reflections of a single underlying phenomenon. Even a small short run predictive power or non zero contemporaneous correlation build up to yield substantial returns predictability at longer horizon if the forecasting variable is persistent. For example, if daily returns are very slightly predictable by a slow-moving (i.e., persistent) variable, that predictability adds up over long horizons. As argued by Cochrane (1999) in a very illuminating way, you can predict that the temperature in Chicago will rise about one-third of a degree per day in spring. This forecast explains very little of the day to day variation in temperature but, because temperature changes are persistent (within each season), it tracks almost all the rise in temperature from January to July. Thus, the  $R^2$  rises with horizon. Precisely, suppose that we forecast excess returns with a forecasting variable  $x$ :

$$\begin{aligned} r_{t+1} &= a + bx_t + e_{t+1} \\ x_{t+1} &= c + \rho x_t + \varepsilon_{t+1} \end{aligned} \tag{2.49}$$

Even for small values of short-horizon  $b$  and  $R^2$  in the first equation above, a large coefficient  $\rho$  in the second equation implies that the long-horizon regression has a large regression coefficient  $b$  and a large  $R^2$ . This regression has a powerful implication: stocks are in many ways like bonds. Any bond investor understands that a string of good past returns that pushes the price up is bad news for subsequent

returns. Many stock investors see a string of good past returns and interpret this as a sign of a bull market, concluding that future stock returns will be good as well. The regression reveals the opposite: a string of good past returns which drives up stock prices is bad news for subsequent stock returns, as it is for bonds.

**Table 2.1**  
**Returns Predictability**

Horizon $k$	$b$	Standard Error	$R^2$
1 years	-1.04	0.33	0.17
2 years	-2.04	0.66	0.26
3 years	-2.84	0.88	0.38
5 years	-6.22	1.24	0.59

*Notes.* This table reports OLS regressions of excess-returns (value-weighted NYSE - Treasury bill rate) on value-weighted price/dividend ratio (reproduced from Cochrane (2001)):

$$r_{t \rightarrow t+k} = a + b (P_t/D_t) + e_t$$

$r_{t \rightarrow t+k}$  denotes the  $k$ th year return. Standard errors use GMM to correct for heteroskedasticity and error autocorrelation.

The co-integrating relation between consumption, asset wealth, and labor income and between consumption and dividends imply that asset prices are set according to the rational valuation formula (RVF), i.e. prices and wealth equal the present value of the rational expectation of future cash flows, either consumption (real cash flows) or dividends, discounted at the equilibrium expected rate of return. The RVF is the solution to a stochastic differential equation where prices equal the present value of the rational expectation of next period dividend or consumption flows and capital gains discounted at the equilibrium expected rate of return. For this differential equation to have a determinate solution, a boundary condition that rules out bubbles must hold. Without this condition (equivalent to requiring that sooner or later any bubble bursts), any self-fulfilling expectation of capital gains would imply a



different yet legitimate solution. In turn, the lack of this boundary condition would imply that the right hand side of (2.47) and (2.48) contains a non-stationary bubble component (in addition to the stationary terms in the rate of future returns and of consumption or dividend growth) and the left hand side would be non-stationary.

The intimate relation between stationarity of the left-hand side of (2.47) and (2.48) and rational valuation has generated intense interest in tests of the co-integrating relation between variables such as prices and dividends or consumption, asset wealth and labor income. While Lettau and Ludvigson (2001) find that consumption, asset wealth and labor income are co-integrated and a large body of evidence suggests that the dividend-price ratio is stationary, see for example Cochrane (1999, 2001), the evidence that prices and dividends are co-integrated is at best weaker. In particular, tests based on the Engle and Granger (1987) methodology find limited evidence of cointegration between dividend and prices, see for example Campbell and Shiller (1987), Diba and Grossman (1988), Froot and Obstfeld (1991), Balke and Wohar (2001). Since prices are much more volatile than dividends, see for example Campbell and Shiller (1987, 1988), it is possible that these tests fail to detect cointegration because the parameters of the cointegrating relation are time varying and, in particular, they display heteroskedastic variability characterized by clustering over time. Heteroskedastic time-variation in the parameters of the cointegrating relation in turn might help explain heteroskedastic excess volatility of prices over fundamentals. Harris, McCabe and Leybourne (2002) introduce a test for stochastic cointegration, where the parameters of the cointegrating relation are allowed to be time varying. This test encompasses the test for cointegration with fixed parameters of the cointegrating relation, defined stationary cointegration. McCabe, Leybourne and Harris (2002) find mixed evidence in favour of stochastic cointegration between stock and dividends but this evidence is stronger than the evidence in favour of stationary cointegration.

### 2.5.6 *The Role of Systematic Skewness*

Non normal return distributions cannot be entirely described by first and second moments. Unless investors display a special type of preferences (quadratic), they care about higher moments. In particular, while NIARA rules out preference for negative portfolio skewness, decreasing absolute risk aversion (DARA) implies preference for positive skewness. As argued by Richter (1960), Levy (1969) and Kraus and Litzenberger (1976), an exact preference ordering for risky portfolios using the first three moments of the portfolio return is possible, in general, only for an investor with a cubic utility function of wealth. Unfortunately, as shown by Levy (1969) and Tsiang (1972), this third degree polynomial utility function is unsuitable to model the preferences of a risk adverse investor. Duly restricted third order Taylor expansions of admissible non-polynomial utility functions can be used instead. Under (NS, RA and) DARA and hence if the investor has a preference for positive portfolio skewness, he should be willing to accept a somewhat lower expected return to hold assets with positive coskewness.

### 2.5.7 *The 3M-CAPM*

Kraus and Litzenberger's (1976) consider the optimal portfolio choice of a representative investor that lives in a 1-period economy. His utility is defined over end of period wealth  $W$ , i.e.  $u = u(W)$ , and it is not restricted to any particular functional form. The only requirement is that it be continuous and three times continuously differentiable over the range of wealth. In this very simple 1-period setting, where the investor does not have to solve the usual optimal consumption-investment decision problem that arises in multi-period (2 or more periods) models, the Euler equation for the maximization of his expected utility is:

$$E_t[u'(W_{m,t+1})r_{i,t+1}] = 0 \quad (2.50)$$



As shown in Appendix A, a third order Taylor expansion of a standardized utility function around the point  $W_0 = E(W) = 1$  yields:

$$\begin{aligned} u(W_{m,t+1}) &\cong [R_{m,t+1} - E(R_{m,t})] + \theta_{1,t} [R_{m,t+1} - E(R_{m,t})]^2 + \theta_{2,t} [R_{m,t+1} - E(R_{m,t})]^3 \\ &\cong [r_{m,t+1} - E(r_{m,t})] + \theta_{1,t} [r_{m,t+1} - E(r_{m,t})]^2 + \theta_{2,t} [r_{m,t+1} - E(r_{m,t})]^3 \end{aligned} \quad (2.51)$$

Here,  $\theta_{1t} = \frac{1}{2}u''(1)$  and  $\theta_{2t} = \frac{1}{6}u'''(1)$ . In the second line of (2.51), I use excess returns instead of returns because, in this simple 1-period setting where the distinction between unconditional and conditional moments is irrelevant, the risk free rate is known with certainty (also conditionally) and, therefore,  $R_{m,t+1} - E_T(R_{m,t}) = r_{m,t+1} - E_T(r_{m,t})$ . Differentiating (2.51) once with respect to wealth, marginal utility can be approximated as follows:

$$u'(r_{m,t+1}) \cong 1 + 2\theta_{1,t} [r_{m,t+1} - E(r_{m,t})] + 3\theta_{2,t} [r_{m,t+1} - E(r_{m,t})]^2 \quad (2.52)$$

Using (2.52) in (2.50) yields Kraus and Litzenberger's (1976) 3M-CAPM. Interpreting marginal utility  $u'(W_{m,t+1})$  as a SDF, (2.50) can be seen as a version of (2.1) where  $x_{t+1} = r_i$  and, because  $r_i$  is an excess return,  $p_t = 0$ . In (2.52), the SDF is approximated as a linear function of the market excess return and its square and thus it can be seen as an instance of (2.14) with  $f_{1,t+1} = r_{m,t+1} - E(r_{m,t})$  and  $f_{2,t+1} = [r_{m,t+1} - E(r_{m,t})]^2$ ,  $a_t = 1$ ,  $b_{1,t} = 2\theta_{1,t}$ ,  $b_{2,t} = 3\theta_{2,t}$ . Applying (2.6), and dropping time-subscripts for notational simplicity, (2.50) and (2.52) imply:

$$E(r_i) = -\frac{\text{Cov}[u'(r_m | \theta), r_i]}{E[u'(r_m | \theta)]} = -\frac{E\left\{ [u'(r_m | \theta) - E(u'(r_m | \theta))] [r_i - E(r_i)] \right\}}{E[u'(r_m | \theta)]} \quad (2.53)$$

Here, differentiating (2.52) once,  $u''(r_m | \theta) = 2\theta_1 + 6\theta_2 r_m$  and, differentiating it once more,  $u'''(r_m | \theta) = 6\theta_2$ . Finally, multiplying and dividing the first and second term on the right-hand side of this equation by, respectively,  $E[r_m - E(r_m)]^2$  and  $E[r_m - E(r_m)]^3$  and re-arranging, we can write<sup>15</sup>:

$$E(r_i) \cong \delta_1 \beta_i + \delta_2 \gamma_i \quad (2.54)$$

Where,

$$\delta_1 = \frac{-E[u''(r_m | \theta)]E[r_m - E(r_m)]^2}{E[u'(r_m | \theta)]} \quad (2.55)$$

$$\delta_2 = \frac{-\frac{1}{2}E[u'''(r_m | \theta)]E[r_m - E(r_m)]^3}{E[u'(r_m | \theta)]} \quad (2.56)$$

$$\beta_i = \frac{E[(r_i - E(r_i))(r_m - E(r_m))]}{E[r_m - E(r_m)]^2} \quad (2.57)$$

$$\gamma_i = \frac{E[(r_i - E(r_i))(r_m - E(r_m))^2]}{E[r_m - E(r_m)]^3} \quad (2.58)$$

Here, the coefficient  $\delta_1$  is the beta premium and the coefficient  $\delta_2$  is the gamma premium. This is a beta-gamma representation of the implications of (2.50) and (2.52) for the cross-section of asset returns. It is different from the beta-pricing representations described in (2.20) because beta and gamma are not multiple regression coefficients. The assumption of greed implies  $E[u'(r_m | \theta)] > 0$  and, under RA,  $E[u''(r_m | \theta)] \leq 0$ . Thus, since  $E[r_m - E(r_m)]^2 \geq 0$ , the beta coefficient  $\delta_1$  is

---

<sup>15</sup> I also assume that the second derivative of the utility function does not depend on the interaction between market and asset unexpected returns,  $Cov\{u''(r_m | \theta), [r_i - E(r_i)][r_m - E(r_m)]\} = 0$ , and that the third derivative does not depend on the interaction between squared market unexpected returns and asset unexpected returns  $Cov\{u'''(r_m | \theta), [r_i - E(r_i)][r_m - E(r_m)]^2\}$ . These are very useful and reasonable simplifications that, intuitively, correspond to the requirement that absolute risk aversion and preference towards skewness do not depend on the relation between a single asset and the market portfolio or its square (rather, they should depend only on the latter, i.e. the market return and its square). Essentially, only changes in overall wealth and in its volatility should determine moves along the utility function and, therefore, changes in the point at which its derivatives are evaluated.



*positive* for risk-averse, greedy investors. If the market portfolio skewness is negative (as it is often the case empirically) and if there is a market reward for holding assets with negative systematic asset coskewness, the gamma coefficient  $\delta_2$  is *positive*. This can also be seen by noting that, under the assumption of greed and NIARA,  $E[u'(r_m | \theta)] > 0$  and  $E[u'''(r_m | \theta)] \geq 0$  respectively. Since empirical market portfolio skewness is usually found to be negative, i.e.  $E[r_m - E(r_m)]^3 \leq 0$ , then  $\delta_2 \geq 0$ . While  $\delta_1$  represents investors' reward for systematic variance, i.e. for holding assets that increase the volatility of the overall market portfolio,  $\delta_2$  compensates investors for systematic negative skewness, i.e. for holding assets that decrease the skewness of the overall market portfolio (that cause the distribution of portfolio returns to be skewed to the left).

### 2.5.8 The 3M-(C)CAPM

While allowing utility to contain a cubic term in wealth, its parameters could be allowed to be time varying. For example, the elements of  $\theta$  in (2.53) could be specified as a function of conditioning information. A particularly interesting possibility is that they vary with the business cycle or that they are a function of conditioning variables that represent investors' expectations about future returns. This would yield a conditional version of Kraus and Litzenberger (1976) 3M-CAPM. Following a similar approach, Harvey and Siddique (2000) propose a conditional asset pricing equation where expected asset excess returns are a function of their conditional covariance and coskewness with the market portfolio and the prices of covariance and coskewness risk also vary over time:

$$\begin{aligned} E_{t-1}(r_{i,t}) &= \xi_t \text{Cov}_t(r_{i,t+1}, m_{t+1}) \\ &= \xi_{1,t} \text{Cov}_t(r_{i,t+1}, r_{m,t+1}) + \xi_{2,t} \text{Cov}_t(r_{i,t+1}, r_{m,t+1}^2) \end{aligned} \quad (2.59)$$

Where:

$$\xi_{1,t} = -b_{1,t} = \frac{\text{Var}_t(r_{m,t+1}^2)E_t(r_{m,t+1}) - \text{Skew}_t(r_{m,t+1})E_t(r_{m,t+1}^2)}{\text{Var}_t(r_{m,t+1})\text{Var}_t(r_{m,t+1}^2) - [\text{Skew}_t(r_{m,t+1})]^2}$$

$$\check{\zeta}_{2,t} = -b_{2,t} = \frac{Var_t(r_{m,t+1})E_t(r_{m,t+1}^2) - Skew_t(r_{m,t+1})E_t(r_{m,t+1})}{Var_t(r_{m,t+1})Var_t(r_{m,t+1}^2) - [Skew_t(r_{m,t+1})]^2}$$

Here, the symbol  $Skew_t(r_{m,t+1}) = E_t\{[r_{m,t+1} - E_t(r_{m,t+1})]^3\}$  represents the skewness of the market portfolio in  $t+1$  conditional on information available at time  $t$ , while the other symbols (e.g. conditional expectation and variance operators) have the usual meaning. The pricing equation in (2.59) can be derived from (2.19) specifying the SDF  $m_{t+1}$  as a quadratic polynomial in the market excess return  $r_{m,t+1}$  with parameters  $a_t$ ,  $b_{1t}$  and  $b_{2t}$  that are allowed to vary over time,  $1 + R_{f,t+1} \cong 1$ , and therefore<sup>16</sup>,  $a_t + b'_{1t}E_t(f_{t+1}) \cong 1$ .

$$m_{t+1} = a_t + b_{1t}r_{m,t+1} + b_{2t}r_{m,t+1}^2 \quad (2.60)$$

Interpreting  $m_{t+1}$  in (2.60) as the SDF implied by a third order Taylor expansion of the representative investor's utility function, the pricing equation in (2.59) can be seen as the cross-sectional implication of a conditional version of the 3 moment CAPM (henceforth, 3M-(C)CAPM). Under this model, if investors like positive portfolio skewness, they should accept a negative risk premium to hold assets with positive coskewness because these assets contribute to increase the skewness of the overall market portfolio. The price of coskewness risk  $\check{\zeta}_{2,t}$ , therefore, should be negative. It is worth at this point highlighting the difference with the 3-moment model derived by Kraus and Litzenberger (1976) where, if market portfolio skewness is negative, positive asset coskewness implies a negative gamma and a positive  $\delta_2$ . In other words, the specification of the systematic third moment premium used by Harvey and Siddique (2000) and by Kraus and Litzenberger (1976) are not equivalent.

---

<sup>16</sup> Recall that, as shown in Section 2.2, in (2.12), and 2.3, in (2.19) and (2.22), without this approximation and the resulting restriction on the relation between the intercept and the mean of the factors the risk free return would show up in the equations for the risk prices.



### 2.5.9 Tests of the 3M-(C)CAPM

Harvey and Siddique (2000) test the 3M-CAPM on Centre for Research on Security Prices (CRSP) NYSE, AMEX and NASDAQ stock data over the period 1963-1993. They find that the 3M-CAPM significantly improves on a 2 moment CAPM specification. They report that coskewness helps explain the cross-section of average excess-returns on 32 industry portfolios and 25 size and book-to-market value sorted portfolios. Moreover, they find that coskewness retains a significant explanatory power even after the inclusion of factors related to size and book to market value that have been found by Fama and French (1992, 1993, 1995) to empirically explain a large portion of the cross-sectional variation in average asset returns. In particular, they find that systematic skewness is important and commands on average a risk premium of 3.6 percent per annum.

Dittmar (2002) specifies a conditional model by expressing the parameters of a quadratic and cubic SDF as linear functions of a set of conditioning variables. The quadratic SDF implies the 3M-(C)CAPM whereas the cubic SDF implies a 4 moment CAPM where preference for co-kurtosis (the systematic fourth moment), is allowed. The conditioning variables include one lag of the market excess-return, of the dividend yield, the spread of the 3 Month T-Bill over the 1 Month T-Bill rate and the 1 Month T-Bill rate itself. He finds evidence of substantial non-linearity in the pricing kernel and that both the quadratic and cubic SDF fit well the cross-section of US industry equity indices average returns over the period 1963-1995. After imposing the regularity conditions on the shape of the utility functions that correspond to *standard* risk aversion, i.e. positive marginal utility, RA and NIARA over all values of wealth, the estimated gamma premium remains statistically and economically significant but it becomes much smaller, thus considerably reducing the ability of the estimated 3 and 4 moment conditional specifications to explain the cross-section of average returns.

Post, Levy and van Vliet (2003) criticise previous empirical tests of the 3M-CAPM, such as Harvey and Siddique (2000), on the grounds that they fail to check whether the decreasing marginal risk-aversion requirement is satisfied by the estimated pricing model. Consistently with Dittmar (2002), they show that the *gamma* (standardised asset co-skewness) premium turns out very small when the appropriate regularity conditions (risk aversion) are imposed on the shape of the investor utility function. In fact, fitting a cubic utility to data on the Fama and French (1995) market portfolio and on 10 size-ranked portfolios for the period 1963-2001, their estimated expected utility function does not satisfy the concavity requirement over the relevant wealth interval and thus the market portfolio is not guaranteed to be efficient for the representative investor. Moreover, they find that the market portfolio is likely to minimize the sample expected utility, rather than maximize it as predicted by the 3M-CAPM.

#### 2.5.10 3M-CAPM vs. (C)CAPM

Even though the (C)CAPM uses the assumption that investors have a quadratic utility function and its pricing kernel does not incorporate 3<sup>rd</sup> order terms, the unconditional implications of both the 3M-CAPM and the (C)CAPM contain a premium for a cross third moment of asset returns. The 3M-CAPM contains a premium for the cross third moment between asset return and the square of the market return, i.e. a premium for  $Cov(r_{i,t}, r_{m,t}^2)$ . The (C)CAPM contains a cross third moment between the asset return, the market return and a conditioning variable that influences marginal utility of market wealth, i.e. a premium for  $Cov(r_{i,t}, z_{t-1}r_{m,t})$ . Equivalently, asset coskewness can be seen as the covariance between the asset return and market volatility<sup>17</sup>, whereas in the (C)CAPM the expression  $Cov(r_{i,t}, z_{t-1}r_{m,t})$  can be interpreted as the covariance between the asset return and the sensitivity of the market return to the conditioning variable. In other words, in the 3M-CAPM investors are rewarded for holding assets that perform poorly at

---

<sup>17</sup> More accurately, coskewness should be seen as the covariance with the realization of the market second moment.



times of high market volatility, whereas in the (C)CAPM they are rewarded for holding assets that do poorly when the return on the stock market portfolio is very reactive to the conditioning variable, and hence when it is very reactive to either returns on non market wealth or expected stock market returns. There are a number of circumstances under which the (C)CAPM expression  $Cov(r_{i,t}, z_{t-1}r_{m,t})$  could proxy for asset coskewness and vice versa. This would be the case if  $z_{t-1}$  was a good proxy for  $r_{m,t}$  and hence when the former forecasts the latter.

#### *2.5.11 The Role of Idiosyncratic Income Risk*

All the (rational) asset pricing models mentioned above predict that expected asset excess-returns can be explained on the basis of their relation to one or more pervasive risk factors. As remarked by Chen, Roll and Ross (1986), financial theory has focussed on systematic influences as the likely sources of risk, assuming the ability of investors to hold diversified portfolios. The general conclusion is that a risk premium is required to compensate for the influence of systematic economic news on the payoff of a particular asset, but no extra reward can be earned by needlessly bearing diversifiable risk. Under this theoretical perspective, therefore, no priced risk premium should be related to the residual variance of stock returns. The recent literature, however, has re-examined the relation between risk and return focussing on the role played by total risk, including idiosyncratic risk.

Models such as the CAPM based on complete capital markets are underpinned by the abstraction of the retired wealthy investor or, alternatively, by the assumption that investors live in a world where all sources of income (including labour income) correspond to traded securities. Malkiel and Xu (2000) show that, under an extended version of the CAPM, when some individuals are not fully diversified, nobody can hold the market portfolio and the relevant measure of risk is a combination of systematic and idiosyncratic variance. Merton (1987) develops a model of capital market equilibrium with rational investors and limited information. In the model, investors only know about a subset of the available securities and thus diversification



is, in general, incomplete. Among the main predictions is that idiosyncratic risk has implications for both the cross-section of asset expected returns and for the market expected return. More generally, if we allow for incomplete capital markets, we must recognise that investors' utility might fluctuate as a result of shocks that do not correspond only to the volatility of their traded assets holdings. This can be modelled within a conditional asset pricing framework by either introducing factors or conditioning variables that correspond to sources of variation of marginal utility unrelated to the stock market.

Even in incomplete markets, however, the risk associated with uninsurable idiosyncratic shocks matters for the pricing of financial assets only if these affect the average investor. A classical example is represented by unemployment income shocks. They are not insurable because of moral hazard and asymmetric information problems (a case of missing markets). This makes it impossible for investors to eliminate these risks through diversification<sup>18</sup>. They also affect the average investor. Virtually everybody is exposed to the risk of becoming unemployed. Investors might try to use traded financial assets to hedge this risk but, in doing so, they cannot take each others' offsetting positions because everybody has an exposure of the same sign. As a consequence, their hedging positions influence asset prices and expected returns. Therefore, while non-insurable shocks are by definition *entirely* non-diversifiable (because of 'missing markets', *both* the average variance and average covariance are not diversifiable, rather than *just* the latter as in the case of insurable shocks), they must affect the average investor and thus, in a sense, they must be "systematic" in order to have asset pricing implications. If, however, idiosyncratic shocks are contemporaneously correlated with systematic stock market shocks, they can be quickly traded away by taking hedging positions in the stock market portfolio and covariance with the stock market portfolio would be again the only relevant risk exposure.

---

<sup>18</sup> If there was no moral hazard and asymmetric information problems, investors would pool these risks together alongside all other risks. Everybody would end up holding a small share of the market portfolio, including everybody else's labour skills, and only the average covariance of labour income with the market portfolio would have to be borne by somebody. The average idiosyncratic variance would be diversified away.



Therefore, while a condition for sources of non-insurable idiosyncratic risk to affect asset prices is that they affect the average investor, they should be uncorrelated with the stock market in order for this type of risk to have asset pricing implications beyond those captured by exposure to the stock market alone. To reproduce these conditions, Constantinides and Duffie (1996) impose that, while idiosyncratic shocks are not correlated<sup>19</sup> with stock market returns and hence they cannot be either diversified or traded away, the distribution of the former across individuals depends on the realization of the latter. A suitable non-linearity in the utility function then ensures that idiosyncratic marginal utility shocks are systematic and correlated with the market return even though idiosyncratic shocks are not correlated with the latter.

### 2.5.12 *The Market Risk Premium in Integrated and Segmented Markets*

To illustrate the effect of the assumptions that markets are integrated and that investors are able to fully diversify, it is useful to compare the two limiting cases of complete market integration and complete market segmentation. Preliminarily, let asset excess-returns be generated by the following model:

$$r_{i,t+1} = E_t(r_{i,t+1}) + \varepsilon_{i,t+1} \quad (2.61)$$

Here,  $\varepsilon_{i,t+1}$  denotes a conditionally zero mean random residual. The aggregate market excess-return equation can be obtained averaging (2.61) across all assets. With  $N$  assets, the aggregate market portfolio excess-return  $r_{m,t+1}$  is the following:

$$r_{m,t+1} = \sum_{i=1}^N w_{i,t} r_{i,t+1} = \sum_{i=1}^N w_{i,t} E_t(r_{i,t+1}) + \sum_{i=1}^N w_{i,t} \varepsilon_{i,t+1} \quad (2.62)$$

---

<sup>19</sup> If idiosyncratic and systematic stock market shocks are contemporaneously correlated covariance with the stock market portfolio would be the only relevant risk exposure.

If assets are priced in integrated markets, the aggregate expected risk-return equation implied by the CAPM is (2.43) with constant  $a_t$  and  $b_t$ :

$$\begin{aligned} E_t(r_{m,t+1}) &= RRA_t \text{Var}_t(r_{m,t+1}) \\ &= RRA_t MKT_t \end{aligned} \quad (2.63)$$

Here,  $MKT_t = \text{Var}_t(r_{m,t+1})$  denotes the conditional variance of the market portfolio conditional on information available at time  $t$  and  $RRA_t = -W_{m,t} \frac{u_t''(W_{m,t})}{u_t'(W_{m,t})}$  is the relative risk aversion coefficient.

Assume now that assets are priced in *segmented* markets so that investors cannot diversify (i.e. each investor can hold only one asset). In this extreme situation, applying (B.13) from Appendix B, the expected risk-return relation for each single asset resembles the risk-return relation for the market portfolio implied by (2.43) under an integrated capital market:

$$E_t(r_{i,t+1}) = RRA_{it} \text{Var}_t(r_{i,t+1}) \quad (2.64)$$

Here,  $RRA_{it}$  is the relative risk aversion coefficient of investor  $i$  that holds asset  $i$ . With  $N$  assets and using (2.62) and (2.64), the aggregate portfolio return  $r_{m,t+1}$  is the following:

$$r_{m,t+1} = \sum_{i=1}^N w_{i,t} RRA_{it} \text{Var}_t(r_{i,t+1}) + \sum_{i=1}^N w_{i,t} \varepsilon_{i,t+1} \quad (2.65)$$

Taking conditional expectations of both sides of (2.65) and assuming that relative risk aversion is constant in wealth, the aggregate risk-return equation is:



$$\begin{aligned}
E_t(r_{m,t+1}) &= \sum_{i=1}^N w_{i,t} CRR A_{it} Var_t(r_{i,t+1}) \\
&= CRR A_t \sum_{i=1}^N w_{i,t} Var_t(r_{i,t+1}) \\
&= CRR A_t \cdot VAR_t
\end{aligned} \tag{2.66}$$

Here,  $CRR A_{it}$  is a constant relative risk aversion coefficient implied by an appropriate utility function,  $CRR A_t = \sum_{i=1}^N w_{i,t} CRR A_{it}$  and  $VAR_t = \sum_{i=1}^N w_{i,t} Var_t(r_{i,t+1})$  is the average total conditional variance of asset returns. The second equality in (2.66) follows from the first one because relative risk aversion is by assumption constant in wealth and thus it is also independent of the second moments of wealth. Comparing (2.66) with (2.63) clarifies that, in segmented capital markets or, more generally, whenever investors do not hold fully diversified portfolios, the market risk premium depends on aggregate total risk, which in turn includes both systematic and idiosyncratic risk. In a more sophisticated setting, this is also one of the main predictions of Merton's (1987) model with limited information and incomplete diversification.

French, Schwert and Stambaugh (1987) found evidence that, while actual volatility and actual returns are negatively correlated, the expected component of the stock market excess return is positively related to the predictable stock market volatility. Both these results provide evidence of a positive relation between the market risk premium and expected market volatility<sup>20</sup>. Yet, many empirical studies fail to agree on the sign of this relation. Estimates of the simple risk-return relation range from significant positive, such as in Harvey (1989), Turner, Startz and Nelson (1989), to significant negative, as in Campbell (1987), Glosten, Jagannathan and Runkle (1993). Whitelaw (1994) finds that the simple lagged market variance-return relation is positive but statistically insignificant. Within Merton's (1973) ICAPM, Scruggs

---

<sup>20</sup> A positive relation between expected components of market return and volatility implies that, when expected volatility increases, it raises the expected market return and actual volatility. Higher expected returns lead to negative actual returns because prices fall. Hence, it implies a negative relation between actual volatility and returns.

(1998) specifies a conditional two-factor model of the market risk premium. In this model the market risk premium is a function of the conditional market variance and the conditional covariance between market excess returns and a variable that describes the state of the investment opportunities in the economy. The state variable chosen by Scruggs (1998) is the conditional excess return on long term US government bonds. Goyal and Santa-Clara (2001, 2003) provide puzzling evidence on the trade-off between risk and return on the US stock market. Using the CRSP database of daily price data starting in July 1962, they compute monthly aggregate excess-return and monthly average total excess-return variance time series for a portfolio that includes all the stocks listed in the main US stock markets (AMEX, NYSE, NASDAQ). They find that there is a significant positive relation between excess-return on the market portfolio and lagged average total variance. Moreover, they find that market variance alone has little explanatory power for market returns. They conclude that there is a significant relation between risk and return, except that risk is measured as total risk, including idiosyncratic risk, rather than only systematic risk. Harvey (1996 and 2000) recognises the link between financial market integration and the relative importance of systematic and total volatility in driving aggregate returns. He finds that the return on little integrated emerging markets is more related to its own total variance than to the variance of the world market portfolio relative to the developed countries markets.

## **2.6. Behavioural Models**

The behavioural finance explanation of the stylized features of the distribution of asset returns also belongs to the new paradigm. While it does not rule out time varying risk and risk premia, it allows for investors' irrationality and market inefficiency. Under this approach, it is admissible that asset prices and expected returns are not the solution to a general equilibrium model with fully rational, risk averse economic agents and competitive financial markets. See, for a review, Barberis and Thaler (2003)). The literature on limits to arbitrage clarified that, in the presence of noise trader risk, risk-averse market participants with short horizons



(finitely lived) might not have the incentive to trade quickly as to exploit all available information even though financial markets are competitive and hence investors are price takers. This is the perspective advocated, among others, by DeLong, Shleifer, Summers and Waldmann (1990a and 1990b) and Shleifer and Vishny (1997). Noise trader risk is the risk that mispricing caused by the net demand of irrational (and hence uninformed) noise traders might worsen in the short run before trades by rational (and hence informed) traders manage to correct it. The relevant notion of rationality is, in this context, the definition embedded in Muth's (1961) rational expectation hypothesis.

The behavioural perspective also allows for non standard utility functions where investors either do not have unambiguously defined preferences over consumption or they display risk seeking over certain portions of the utility function domain. For example, Prospect Theory and Cumulative Prospect Theory, formulated by Kahneman and Tversky (1979) and Tversky and Kahneman (1992) respectively, imply framing and S-shaped utility functions defined over gains and losses instead of over consumption and wealth as in the standard Expected Utility framework. In particular, these utility functions display risk aversion over gains and risk seeking over losses below a threshold. Behavioural Portfolio Theory, advocated by Shefrin and Statman (2002), predicts instead risk aversion over losses and risk seeking over gains and thus an inverse S-shaped utility function. These non standard utility functions rationalize evidence that investors, under certain circumstances, display risk seeking behaviour. Active stock traders appear to play negative-sum games and their behavior can sometimes be interpreted as 'gambling' (see Statman (2002)). In addition, psychologists led by Kahneman and Tversky (1979) find experimental evidence for local risk seeking behavior. More specifically, Post and Levy (2002) argue that a number of celebrated asset pricing anomalies, such as the low average yield on stocks with large capitalization, growth stocks and past winners, could be explained by risk aversion over losses and risk seeking over gains.



Numerous contributions from the literature on non standard utility theory and behavioural asset pricing (see for a review Shefrin (2005)), thus, admit a non linear pricing kernel that implies non concavity of the utility function over certain ranges of wealth, and thus an increasing SDF and a violation of RA. Friedman and Savage (1948) and Markowitz (1952) argue that the willingness to purchase both insurance and lottery tickets implies that marginal utility is increasing over a range. See Hartley and Farrell (2002) and Post and Levy (2002) for a recent discussion. S-shaped utility functions, such as the function implied by Kahneman and Tversky's (1979) prospect theory, do not satisfy either RA or NIARA. Inverse S-shaped utility functions, such as the specification implied by Shefrin and Statman's (2002) behavioural portfolio theory, violate RA but satisfy NIARA at every point in the domain where the function is differentiable.

## **2.7. Absolute vs. Relative Pricing**

Financial theory has extensively addressed the issue of how to model the mean behaviour of asset returns and link it to other variables. In particular, asset pricing models relate mean returns to higher moments. The latter are typically cross-moments formed between the asset return and other variables. These variables can be either economic fundamentals and other non-asset variables, or returns on other assets such as the market portfolio. The first approach is known as absolute pricing, e.g. Lucas' (1978) Consumption CAPM, whereas the second is known as relative pricing, e.g. the Sharpe (1964) and Lintner (1965) CAPM and especially the Asset Pricing Theory (APT), proposed by Ross (1976). The APT requires for its derivation less restrictive assumptions than the CAPM, such as that investors are greedy, that markets are frictionless (or, at least, that diversification is not too costly) and that the returns variance-covariance matrix has a well defined factor structure. The latter requirement guarantees that diversified portfolios can be closely replicated by portfolios that mimic the exposure to single factors. It does not, however, require market completeness (or, equivalently, the representative investor assumption). It provides a no-arbitrage pricing relation between diversified portfolios of assets



based on their sensitivity to a set of pervasive risk factors and on the equilibrium risk premium for the exposure to each factor. Chen, Roll and Ross (1986) proxy for the factors using macro-economic variables deemed to drive the variation in stock returns. Within a multi-factor model for asset returns derived from the APT, Koutulas and Kryzanowski (1996) estimated conditional time-varying risk premia and conditional volatilities associated with each pervasive risk factor. They found that five pervasive risk factors, namely the lag of industrial production, the Canadian Index of 10 Leading Indicators, the US Composite Index of 12 leading Indicators, the exchange rate and the residual market factor, have priced risk premia, including the residual market factor.

For large and diversified portfolios, the implications of the CAPM and of the APT are the same when there is only one pervasive risk factor, the market portfolio. In this case, the expected excess return on the market portfolio would be the only risk premium priced in equilibrium for any diversified portfolio. Neither the CAPM nor the APT admit any risk premium related to idiosyncratic risk (exposure to asset-specific, non-pervasive risk factors), which is expected to be diversified away. It should be noted however that the CAPM, contrary to a popular misinterpretation, is not a special case of the APT. The latter imposes an assumption, namely that the idiosyncratic residuals are uncorrelated, that the CAPM does not require. In the CAPM, idiosyncratic residuals are uncorrelated only on average (with capitalization weights). This is not an assumption, but an implication that follows by construction from the CAPM prediction that these residuals are the error terms of the regression of a set of asset excess returns on their own capitalization-weighted average, namely the market excess returns.

## **2.8. Summary and Conclusions**

In this section I have summarized the important developments in asset pricing theory along with the transition from the old to the new paradigm of asset returns, and I have shown how the various asset pricing models can be seen as specializations of

the general SDF model. I have thus reviewed a number of specifications of this model. Whenever possible, I highlighted the connections between the implications of the various asset pricing models and interesting patterns in equity returns and their moments. I also discussed the role of volatility in asset pricing theory and the assumptions implied by empirical tests of the conditional risk-return relation.

The SDF representation of the asset pricing problem is surprisingly flexible, yet it allows explanations for the observed patterns in asset returns to be generated in a rigorous and testable manner. The only requirement is that the SDF be linear in the factors. Since this approach allows for considerable flexibility in specifying the functional form of the SDF, it can capture non-linearity in the behaviour of marginal utility and time variation in the parameters of the utility function. It therefore serves as a useful framework to specify alternative asset pricing models that allow for a variety of factors to be priced in the time series and cross section of asset returns under alternative assumptions about the multivariate distribution of asset returns, investors' preferences and market completeness. All the asset pricing models estimated in Chapter 4, 5 and 7 can be seen as specializations of a SDF model. Chapter 4 will explicitly derive a beta-pricing representation of the 3M-CAPM from a quadratic SDF model. Chapter 5 will test whether aggregate idiosyncratic risk is priced in the time series of aggregate returns and thus whether it is a candidate to enter the SDF equation as a factor. In Chapter 7, I will allow for heteroskedasticity and serial dependence in the second moments of the market return risk factor.



## Chapter 3: The Second Moments Literature

### 3.1. Introduction

A rich literature on the second moments (volatilities and correlations) of the empirical distribution of asset returns both contributed to and was promoted by the abandonment of the *i.i.d.* hypothesis. In this section, I will review a number of contributions on empirical models and methodologies for the estimation of second moments of asset returns. In doing so, I will pay special attention to clarify the differences between conditional and unconditional estimates.

I will first review, in Section 3.2, the literature on market-wide volatility and the more recent contributions on individual stock volatility, idiosyncratic stock volatility and stock correlations. I will introduce, in Section 3.3, the variance decomposition methodology proposed by Campbell, Lettau, Malkiel and Xu (2001), henceforth CLMX (2001), because it provides a neat analytical framework to isolate the main components of the variability of the typical asset. I will then examine the main problems that arise when measuring variance components within a CAPM framework. In doing this, I will pay special attention to the instability of the beta coefficients in the market model. In Section 3.4, I will outline the distinction between unconditional and conditional second moments estimation methodologies and I will introduce the literature on multivariate conditional second moments. I will then discuss at some length Engle's (2002) and Engle and Sheppard's (2001 and 2002) dynamic conditional correlation GARCH (DCC-GARCH) model because it provides a feasible way to estimate the parameters that govern the dynamics of large variance-covariance matrices and, in particular, of the associated correlation process. In the final section, I will draw together the main conclusions of the chapter.

### 3.2. Second Moments

As a preliminary definitional matter, recall that variance and correlation are second moments centred around the mean, i.e. the first moment, of a random variable. When the random variables under consideration are returns, volatility is their standard deviation, i.e. the square root of variance and correlation is the covariance between returns standardised by their own volatility. These variables will generally be the focus of the discussion that follows. Developments in financial time series econometrics have led to vast improvements in our understanding of the behavior of the second moments of return distributions over time. Early contributions, such as the pioneering work of Officer (1973) and Schwert (1989), popularized the notion that stock market volatility changes over time. It therefore became clear that the assumption that returns at different points in time were drawn from the same conditional distribution was too restrictive. This represented a lethal blow for the *i.i.d.* hypothesis. Once it became clear that asset volatility and, more generally, second moments, were time-varying, researchers became interested in whether and how they could be modelled.

#### 3.2.1 *Systematic vs. Idiosyncratic Volatility*

A striking feature of the extant literature on financial volatility is the overwhelming prevalence of contributions on aggregate market risk. Partial surveys of this enormous literature are given by Bollersev, Chou and Kroner (1992), Ghysel, Harvey and Renault (1996) and Campbell, Lo and MacKinlay (1997). Aggregate market volatility is relevant to any holder of diversified portfolios and to any model of asset returns developed under the general framework of the Sharpe (1964) and Lintner's (1965) CAPM. Under this theoretical perspective, firm volatility is interpreted as idiosyncratic risk that can be diversified away and that therefore deserves no attention. More recently, however, financial researchers have begun to re-examine the nature and the behaviour of risk in equity markets, addressing both market risk and idiosyncratic risk and the closely related issue of the correlation



among asset returns. Recent evidence, provided by among others Barber and Odean (2000) and Benartzi and Thaler (2001), suggest that investors often hold undiversified portfolios, and even if they are keen to diversify, they tend to hold a limited number of assets to reduce transaction costs. Therefore their relevant measure of risk, as shown by Malkiel and Xu (2000), may well be total risk. Barberis and Thaler (2003) provide a review of this ‘insufficient diversification’ puzzle.

In this vein, CLMX (2001) analyse the long-term trends in both firm-level and market volatility in United States stock markets over the period from 1962 to 1997. Using daily data on all the stocks traded throughout the period on three US markets (AMEX, NASDAQ and the NYSE), they show that while market volatility has not exhibited any significant trend, a decline in overall market correlations has been accompanied by a parallel increase in average firm-level volatility. In explaining their findings, CLMX (2001) suggest that they might emanate from a number of factors, such as the tendency for firms to access the stock market earlier in their development, the existence of time varying betas, executive compensation schemes that reward greater stock volatility, and/or the tendency for large conglomerates to be broken into smaller, less diversified corporations. Whatever the cause, the findings of CLMX (2001) have important implications for portfolio management because they impact significantly on the extent to which diversification can reduce idiosyncratic risk. A conventional rule of thumb, based on Bloomfield, Leftwich and Long (1977), suggests that a randomly chosen portfolio of 20 stocks produces most of the reduction in idiosyncratic risk that can be achieved through diversification. The CLMX (2001) finding that average firm-level volatility has trended upwards in United States stock markets implies that a growing number of stocks is needed to achieve any desired level of diversification. On the other hand, the lower average correlation among the stock returns has increased the potential benefit from diversification because it implies a smaller contribution to portfolio risk of the portion that cannot be diversified away.

### 3.2.2 *Clustering*

An empirical feature of volatility and, more recently, correlation that has attracted considerable attention is that the time series of their realizations tend to exhibit a clustering behaviour. There are periods of high volatility and high correlations and periods when asset returns tend to be more stable and less correlated. Therefore, returns series often display excess kurtosis relative to the multivariate normal distribution (see, for example, Gallant, Rossi and Tauchen (1992)). Mandelbrot (1963) first noticed this phenomenon and French, Schwert and Stambaugh (1987) and Schwert (1989) were among the first to systematically study the clustering behaviour of stock market volatility series. It soon became clear that not only the multivariate distribution of asset returns exhibits time varying second moments but, especially in high frequency series (monthly or higher), the magnitude of their variation is related to how much they change in nearby periods. In other words, volatility and, more generally, the second moments of asset returns display a slow-moving persistent behaviour. After a shock, they tend to mean revert to their long run average rather slowly. It therefore takes time for the effect of each shock to die out. The literature on conditional heteroskedasticity models, initiated by Engle (1982), elegantly captures this behaviour. In these models, conditional second moments depend non-trivially on past states of the world. This conditional heteroskedasticity implies leptokurtosis and therefore an underlying unconditional distribution with fatter tails than under the homoschedasticity hypothesis and the multivariate normal case.

### 3.2.3 *Asymmetry*

Schwert (1989) also noticed that stock market volatility tends to rise during market downturns and to fall during market upturns. The relation between individual stock returns and volatility exhibits a similar pattern (see, among others Cheung and Ng (1992)). Also, while volatility tends to increase after large returns of both positive



and negative sign because of its persistence, it rises more following negative returns than following equally large positive returns. Thus, volatility is usually negatively correlated (again, especially in monthly or higher frequency data) with both contemporaneous and lagged returns. Campbell and Hentschel (1992) labelled the negative correlation between volatility and contemporaneous and lagged returns as contemporaneous and predictive asymmetry, respectively, perhaps because it implies that the distribution of asset returns is skewed to the left.

Two explanations for this phenomenon are popular in the financial literature, the 'leverage effect' and the 'volatility feed back effect'. According to the leverage effect, a large negative return reduces the value of the firm's equity and thus increases financial leverage, in turn rising equity return volatility (e.g. Black (1976) and Christie (1982)) for a given level of asset volatility. More specifically, suppose that bad news regarding operating margins reduce the market value of a firm's assets. The lower asset value must be matched by a decline in the liabilities' value. However, the equity value declines more than the debt value because the latter is more senior. This increases leverage and thus the volatility of the return on equity. Empirically, this effect implies that future volatility is negatively correlated with current returns, and it might also generate contemporaneous negative correlation in low frequency data because of time aggregation. Black (1976) realized, however, that the financial leverage effect alone is empirically insufficient to explain the size of the observed asymmetry. This has also been documented by Christie (1982) and Schwert (1989). Alternatively, if the market risk premium is an increasing function of expected volatility, as implied by many asset pricing models with conditionally time varying moments (such as Merton's (1973) ICAPM) and as suggested by the findings of French, Schwert and Stambaugh's (1987), large negative returns increase future volatility more than positive returns due to a volatility feedback effect (e.g. Campbell and Hentschel (1992)). According to this effect, the impact of negative news on volatility is larger than the effect of positive ones because the former is compounded by the dependence of the expected stock return on expected volatility whereas the latter is dampened. More specifically, both negative and positive news



cause volatility and, because the latter is persistent, increase expected volatility. When the latter rises, however, the expected return also rises, the stock price drops and the realized return is negative. The latter amplifies the negative return initially produced by a piece of bad news, and it dampens the positive return first induced by a piece of good news. Empirically, this effect implies that future returns are positively correlated with current volatility, and future volatility is negatively correlated with current returns (however, time aggregation might also lead to contemporaneous negative correlation in low frequency data).

Bekaert and Wu (2000) provide a unified framework to examine which of these competing explanations is best able to capture asymmetry in equity return volatility. Using data on stocks included in the Nikkei 225 index, they construct a proxy for the Japanese market portfolio and other portfolios with different leverage. They find that although volatility asymmetry is generally present and significant, its source differs across portfolios. More crucially, while it is important to include leverage ratios in the volatility dynamics, their economic effects are mostly dwarfed by the volatility feedback mechanism. They do not find significant asymmetries in conditional betas.

It has also been known for some time that equity return correlations tend to decline in bull markets and to rise in bear markets (De Santis and Gerard (1997), Ang and Bekaert (2002), and Longin and Solnik (2001)). In particular, Longin and Solnik (2001) use extreme value theory to show that it is not volatility per se that affects correlations<sup>21</sup>, but rather the market trend (whether positive or negative returns prevail). Forbes and Rigobon (2002), however, warn against the danger of overestimating the rise in correlations at times when the market in which one of the assets is traded is unusually volatile. They show that the standard correlation coefficient computed using only large absolute returns is higher than the correlation

---

<sup>21</sup> They test whether the correlation of the absolute value of returns in excess of a given threshold go asymptotically to zero as it should under the null of multivariate normality. They find that this is the case only for correlations amongst large positive returns but not for correlations amongst very negative returns. Therefore, returns are multivariate normal in the upper tail but depart from normality in the lower tail.



coefficient computed using all the returns in the sample even when the correlation parameter of the underlying data generating process is the same. They argue that this bias is introduced by estimating correlations conditioning on large absolute deviations, and they show how to adjust the correlation coefficient estimate in order to alleviate the problem. Ang and Chen (2002) discuss the problem of conditioning bias introduced by estimating correlations conditional on high or low returns and on high or low volatility, and they propose a statistic to test the null of multivariate normality against the alternative of asymmetric correlations in downside and upside markets. Recent evidence, provided (among others) by Capiello, Engle and Sheppard (2003), suggests that correlations display asymmetric reactions to *past joint* negative and positive returns. Similarly, Bekaert and Wu (2000) find that volatility feedback is enhanced by a phenomenon that they term covariance asymmetry, i.e. conditional covariances with the market increase significantly only following negative market news.

As noticed by Capiello, Engle and Sheppard (2003), little theoretical framework is available to explain this evidence. Focusing on the asymmetric reaction of correlations to past returns innovations of the same sign (returns are either both positive or both negative), they propose a possible explanation using the notion that risk premia are time-varying. In particular, consider risk premia that vary as a function of time-varying conditional variances. Following negative news on two assets, both their volatilities are likely to increase due to either the leverage or the volatility-feed-back effect. If this increase in volatility feeds into volatility expectations, as would be justified by the persistence of volatility series and as explicitly modelled by the volatility feed-back effect of Campbell and Hentschel (1992), investors demand a higher expected return to hold the two assets, thus requiring their prices to drop further. This implies an increase in their correlation.

### 3.2.4 *Skewness*

Asymmetry in the multivariate distribution of stocks returns encompasses all the phenomena of changing second moments depending on the sign of either current or past realizations of moments of odd order. For example, rising volatility in a market downturn implies a stock return distribution skewed to the left, just like rising correlations following joint negative return realizations imply, *ceteris paribus*, a multivariate distribution skewed to the left.

From this perspective, a more general explanation for the observed asymmetry in the distribution of individual stocks and portfolios of stocks is that variability and dependence in returns is higher and possibly non-linear for large negative returns. In other words, the degree of variability of asset returns and of their co-dependence might be both an increasing function of the absolute size (to explain volatility and correlation time-clustering) and a decreasing function of the size of asset returns. This function might be either linear or non-linear (Patton (2002)). In both cases, volatility and correlation would be linear approximations to the true function that describes the variability and co-dependence of asset returns, locally-valid in a neighbourhood of the current returns.

Patton (2002a) and Patton (2002b), among others, show that asymmetry in the dependence structure, and in particular in the dynamic behaviour of correlations, leads to nonzero asset coskewness and portfolio skewness. Depending on the composition of the portfolio, this behaviour of correlations might imply negative portfolio skewness even if the skewness of the individual assets is on average non-negative or even positive. For example, the skewness of a portfolio of two assets is a function of the skewness of the individual assets, and two co-skewness terms. If the co-skewness terms are negative enough, they might more than offset positive skewness terms. In a portfolio with a large number of assets, the skewness terms



would be diversified away but not the co-skewness terms. If asset co-skewness is on average negative, portfolio skewness would be negative.

The magnitude of both  $Cov_t(r_{it+1}, r_{pt+1}^2)$  and  $Cov(r_{it+1}r_{pt+1}, r_{pt+1})$  is related to the amount of asset coskewness<sup>22</sup>. In particular, the larger the asset coskewness, the larger the two quantities  $Cov_t(r_{it+1}, r_{pt+1}^2)$  and  $Cov(r_{it+1}r_{pt+1}, r_{pt+1})$ . This is true also for portfolio skewness, i.e. for  $Cov(r_{pt+1}, r_{pt+1}^2)$  and  $Cov(r_{pt+1}^2, r_{pt+1})$ . While the first quantity, i.e.  $Cov_t(r_{it+1}, r_{pt+1}^2)$ , has the most direct asset pricing implications and it is therefore popular in the asset pricing literature, see for example the discussion of the 3M-CAPM in Chapter 2, the second quantity, i.e.  $Cov(r_{it+1}r_{pt+1}, r_{pt+1})$ , has the most intuitive implications for the shape of the multivariate distribution of asset returns. It

<sup>22</sup> Recall that asset coskewness with a portfolio is defined in this thesis as  $E_t(\varepsilon_{i,t+1}\varepsilon_{p,t+1}^2)$  whereas portfolio skewness is  $E_t(\varepsilon_{p,t+1}^3)$ , where  $\varepsilon_{i,t+1}$  and  $\varepsilon_{p,t+1}$  are conditionally zero mean return innovations on asset  $i$  and a portfolio, respectively. To see why the larger coskewness, the larger  $Cov_t(r_{it+1}, r_{pt+1}^2)$ , we might rewrite the latter as follows:

$$\begin{aligned} Cov_t(r_{it+1}, r_{pt+1}^2) &= E_t(r_{it+1}r_{pt+1}^2) - E_t(r_{it+1})E_t(r_{pt+1}^2) = E_t\{[\varepsilon_{it+1} + E_t(r_{it+1})][\varepsilon_{pt+1}^2 + E_t(r_{pt+1}^2)]\} - E_t(r_{it+1})E_t(r_{pt+1}^2) \\ &= E_t[\varepsilon_{it+1}\varepsilon_{pt+1}^2 + E_t(r_{it+1})\varepsilon_{pt+1}^2 + \varepsilon_{it+1}E_t(r_{pt+1}^2) + E_t(r_{it+1})E_t(r_{pt+1}^2)] - E_t(r_{it+1})E_t(r_{pt+1}^2) \\ &= E_t(\varepsilon_{it+1}\varepsilon_{pt+1}^2) + E_t(r_{it+1})E_t(\varepsilon_{pt+1}^2) + E_t(\varepsilon_{it+1})E_t(r_{pt+1}^2) + E_t(r_{it+1})E_t(r_{pt+1}^2) - E_t(r_{it+1})E_t(r_{pt+1}^2) \\ &= E_t(\varepsilon_{it+1}\varepsilon_{pt+1}^2) + E_t(r_{it+1})E_t(\varepsilon_{pt+1}^2) = E_t(\varepsilon_{it+1}\varepsilon_{pt+1}^2) + E_t(r_{it+1})\sigma_{p,t}^2 \end{aligned}$$

Similarly, to show that the larger coskewness, the larger  $Cov_t(r_{it+1}r_{pt+1}, r_{pt+1})$ , we might rewrite the latter as follows:

$$\begin{aligned} Cov_t(r_{it+1}r_{pt+1}, r_{pt+1}) &= E_t(r_{it+1}r_{pt+1}^2) - E_t(r_{it+1}r_{pt+1})E_t(r_{pt+1}) \\ &= E_t\{[\varepsilon_{it+1} + E_t(r_{it+1})][\varepsilon_{pt+1}^2 + E_t(r_{pt+1})^2 + 2\varepsilon_{pt+1}E_t(r_{pt+1})]\} - E_t(r_{it+1}r_{pt+1})E_t(r_{pt+1}) \\ &= E_t\{[\varepsilon_{it+1}\varepsilon_{pt+1}^2 + \varepsilon_{it+1}E_t(r_{pt+1})^2 + 2\varepsilon_{it+1}\varepsilon_{pt+1}E_t(r_{pt+1})]\} + \\ &+ [E_t(r_{it+1})\varepsilon_{pt+1}^2 + E_t(r_{it+1})E_t(r_{pt+1})^2 + 2E_t(r_{it+1})\varepsilon_{pt+1}E_t(r_{pt+1})] - E_t(r_{it+1})E_t(r_{pt+1}^2) \\ &= E_t(\varepsilon_{it+1}\varepsilon_{pt+1}^2) + E_t(\varepsilon_{it+1})E_t(r_{pt+1})^2 + 2E_t(\varepsilon_{it+1}\varepsilon_{pt+1})E_t(r_{pt+1}) + \\ &+ E_t(r_{it+1})E_t(\varepsilon_{pt+1}^2) + E_t(r_{it+1})E_t(r_{pt+1})^2 + 2E_t(r_{it+1})E_t(r_{pt+1})^2 - E_t(r_{it+1})E_t(r_{pt+1}^2) \\ &= E_t(\varepsilon_{it+1}\varepsilon_{pt+1}^2) + 2E_t(\varepsilon_{it+1}\varepsilon_{pt+1})E_t(r_{pt+1}) + E_t(r_{it+1})[\sigma_{p,t}^2 + 3E_t(r_{pt+1})^2 - E_t(r_{pt+1}^2)] \\ &= E_t(\varepsilon_{it+1}\varepsilon_{pt+1}^2) + 2E_t(\varepsilon_{it+1}\varepsilon_{pt+1})E_t(r_{pt+1}) + E_t(r_{it+1})[\sigma_{p,t}^2 + 3E_t(r_{pt+1})^2 - \sigma_{p,t}^2 - E_t(r_{pt+1})^2] \\ &= E_t(\varepsilon_{it+1}\varepsilon_{pt+1}^2) + 2E_t(\varepsilon_{it+1}\varepsilon_{pt+1})E_t(r_{pt+1}) + 2E_t(r_{it+1})E_t(r_{pt+1})^2 \end{aligned}$$

implies that adding assets with negative coskewness increases the average covariance in the left-hand tail of the multivariate distribution of assets returns. As a result, the left-hand tail of the portfolio distribution becomes fatter relative to the right-hand tail. This is reflected in a portfolio distribution more skewed to the left.

Recent contributions have examined the distribution of second moments, treated as observable variables, in order to depict a more accurate picture of the multivariate distribution of asset returns. For example, Andersen, Bollerslev, Diebold and Ebens (2001), study realized daily equity return volatilities and correlations obtained from high-frequency intraday transaction prices on individual stocks in the Dow Jones Industrial Average. They find that the unconditional distribution of realized volatilities and correlations implies that the unconditional distributions of realized variances and covariances are highly *right*-skewed. This in turn implies a *left*-skewed multivariate returns distribution<sup>23</sup>.

Increasing volatilities during market downturns necessitates investing in a larger number of stocks to diversify away the idiosyncratic volatility of the typical stock to any desired extent, whereas increasing correlations reduce the effectiveness of diversification strategies precisely at the time when portfolio managers are most reliant on them to reduce the overall risk of their investments. Awareness of the asymmetric behaviour of volatilities and correlations in bull and bear markets can thus lead portfolio managers to better define their asset allocation strategies. Risk managers can also use this information to specify more accurate Value at risk (VaR) models. In particular, incorporating asymmetric correlation behaviour in the generating stochastic process of portfolio returns leads to VaR estimates consistent with asymmetric portfolio returns distributions. This represents a substantial improvement on simpler models that use a multivariate normal distribution<sup>24</sup>.

---

<sup>23</sup> However, they confirm the known result that distributions of the returns scaled by realized standard deviations are approximately Gaussian, implying that the multivariate distribution of conditionally standardized returns is symmetric.

<sup>24</sup> An exhaustive review of the main shortcomings of the traditional VaR model is provided by Szego (2002).



### 3.2.5 *Contagion*

To some extent and especially in an international setting, asymmetry and fat tails in the multivariate distribution of returns can be seen as instances of contagion. King and Wadhvani (1990) define the latter as a significant change in the return correlation across markets. Using daily data, they find support for the presence of contagion between stock markets in the US, the United Kingdom and Japan in the aftermath of the October 1987 US stock market crash. Lee and Kim (1993) also found evidence of contagion. The average cross-market correlation between 12 major international stock markets jumps from 0.23 to 0.39 in the aftermath of the crash.

Baig and Goldfajn (1999) and Forbes and Rigobon (2002) refine the definition of financial contagion as a significant increase in cross-market linkages after a shock to one country or group of countries. Their definition emphasizes the possible emergence of contagion after a major financial crisis hence they focus only on “the crisis period”. Forbes and Rigobon (2002) further pointed out that tests of contagion have to take into account the presence of heteroskedasticity in stock returns. When returns are heteroskedastic, tests of parameter stability based on correlation coefficients are biased. Once this is taken into account and hence tests are adjusted for the presence of heteroskedasticity, the presence of mean contagion is rejected. Correcting for heteroskedasticity in stock returns, Forbes and Rigobon (2002) could not find a significant increase in stock return correlation across stock markets; they concluded there was no contagion but ongoing increased interdependence across markets. While the more influential contributions in the literature focused on correlation of returns across stock markets and thus on spillovers in mean returns, there are also studies that analyzed spillovers of volatility across markets as a form of financial contagion. Spillovers in mean returns amount to an increase of return correlation at times of financial crisis and thus conditional and unconditional excess skewness (relative to the multivariate normal distribution), whereas spillovers of

volatility imply an increase of volatility correlation and thus conditional excess kurtosis and unconditional excess skewness.

Both GARCH and VAR frameworks have been used to estimate the variance-covariance transmission mechanisms between countries. Engle, Ito, and Lin (1990) applied both GARCH and VAR models to test for spillovers in daily exchange rate volatility across Japanese and American foreign exchange markets. They found support for the hypothesis that there are intra-daily volatility spillovers from one foreign exchange market to the other, i.e. they find evidence in favor of the “meteor shower” rather than the “heat waves” hypothesis. Using the GARCH framework, Hamao, Masulis and Ng (1990) analyze short-run price volatility spillovers across London, New York and Tokyo stock markets around the 1987 U.S. market crash. They find evidence of price volatility spillovers from New York to London, from New York to Tokyo and from London to Tokyo but not in other directions. Edwards (2000), on the other hand, estimated a GARCH model of interest rate volatility and found evidence in support of contagion effects from Mexico to Argentina but not to Chile. He interprets this result as an evidence of the curtailing effect of capital controls on volatility contagion in fixed income securities. Applying univariate and bivariate switching volatility models to weekly stock returns for a group of Latin American countries, Edwards and Susmel (2001) find strong evidence of volatility co-movements across countries, especially among the Mercosur countries, that they interpret as evidence of contagion. In addition, they show that high-volatility episodes are, in general, short-lived, lasting from 2 to 12 weeks. Diebold and Yilmaz (2005), in a multivariate VAR framework, study spillovers in both mean return and volatility between 16 major International stock markets, including 12 emerging markets. They report that return spillovers display an upward trend, consistent with increasing interdependence and financial integration, but no contagion. They find however strong evidence of contagion in volatility spillovers.



### 3.3. Variance Decomposition

CLMX (2001) proposed a useful methodology to decompose the total variance of the average stock into its market and idiosyncratic components. I begin by writing a simple linear model of asset excess returns:

$$r_{i,t} = \beta_i r_{m,t} + \varepsilon_{i,t} = r_{m,t} + \eta_{i,t} \quad (3.1)$$

Here,  $r_{m,t}$  is the excess return on the market portfolio of  $N$  assets,  $r_{i,t}$  is the excess return on the asset  $i$  at time  $t$ ,  $i = 1, \dots, N$ ,  $\beta_i$  is the coefficient of the regression of  $r_{i,t}$  on  $r_{m,t}$  with no constant term,  $\varepsilon_{i,t}$  is a regression idiosyncratic<sup>25</sup> residual. Finally,  $\eta_{i,t}$  is a market-adjusted excess return on the asset  $i$  computed according to the second equality in (3.1). In this representation, returns on the assets included in the market portfolio  $m$  are the sum of a linear function of the return on the portfolio and of an error term. Alternatively, they are written as the sum of a portfolio component and of a portfolio-adjusted component. Letting  $w_{i,t}$  denote the weight of asset  $i$ , we can compute the weighted average of the variance of the returns on the  $N$  assets in the market portfolio:

$$\sum_{i=1}^N w_{i,t} \text{Var}(r_{i,t}) = \text{Var}(r_{m,t}) + \sum_{i=1}^N w_{i,t} \text{Var}(\eta_{i,t}) + \sum_{i=1}^N w_{i,t} 2\text{Cov}(r_{m,t}, \eta_{i,t}) \quad (3.2)$$

From (3.1),  $\eta_{i,t} = \varepsilon_{i,t} + r_{m,t}(\beta_i - 1)$ . Thus the weighted average of the variance of the returns becomes:

$$\sum_{i=1}^N w_{i,t} \text{Var}(r_{i,t}) = \text{Var}(r_{m,t}) + \sum_{i=1}^N w_{i,t} \text{Var}(\eta_{i,t}) + 2\text{Var}(r_{m,t}) \sum_{i=1}^N w_{i,t} (\beta_{i,m} - 1) \quad (3.3)$$

---

<sup>25</sup> Notice that each  $\varepsilon_{it}$  and  $\varepsilon_{jt}$ , with  $i \neq j$ , are not in general independent. However, they are uncorrelated on average by construction.

Recalling that the weighed average of the  $\beta_i$  coefficients is equal to 1, the last term on the right collapses to zero, and we are left with the following portfolio variance decomposition:

$$\begin{aligned}
 VAR_t &= \sum_{i=1}^N w_{i,t} Var(r_{i,t}) \\
 &= Var(r_{m,t}) + \sum_{i=1}^N w_{i,t} Var(\eta_{i,t}) \\
 &= Var(r_{m,t}) + IDIO_t
 \end{aligned} \tag{3.4}$$

This decomposes the average excess return variance ( $VAR_t$ ) across all assets in the market portfolio into two components; the variance of the excess return on the market portfolio ( $MKT_t$ ) and an average idiosyncratic component ( $IDIO_t$ ). To explain what  $IDIO_t$  represents, I rewrite it as follows:

$$\begin{aligned}
 \sum_{i=1}^N w_{i,t} Var(\eta_{i,t}) &= \sum_{i=1}^N w_{i,t} Var(r_{i,t} - r_{m,t}) = \sum_{i=1}^N w_{i,t} Var(\beta_i r_{m,t} + \varepsilon_{i,t} - r_{m,t}) \\
 &= \sum_{i=1}^N w_{i,t} [Var(\varepsilon_{i,t}) + (\beta_i - 1)^2 Var(r_{m,t})] \\
 &= \sum_{i=1}^N w_{i,t} Var(\varepsilon_{i,t}) + Var(r_{m,t}) \sum_{i=1}^N w_{i,t} [(\beta_i - 1)^2] \\
 &= \sum_{i=1}^N w_{i,t} Var(\varepsilon_{i,t}) + Var(r_{m,t}) CSV(\beta_i)
 \end{aligned} \tag{3.5}$$

Here,  $CSV(\beta_i) = \sum_{i=1}^N w_{i,t} [(\beta_i - 1)^2]$  is the cross-sectional variance of  $\beta_i$ . As shown by CLMX (2001),  $CSV(\beta_i)$  is relatively small and not very volatile. Therefore,  $IDIO_t$  has the important property of being approximately equal to the average variance of the residuals from the regression of asset excess returns on the market portfolio. Since  $r_{m,t}$  is the excess-return on the market portfolio, the first equality in (3.1) is seen as an empirical version of the Sharpe (1964) and Lintner (1965) security market line (SML) and  $\varepsilon_{i,t}$  is the usual CAPM idiosyncratic residual. However, while (3.1)



contains a term in the asset beta, betas do not appear in (3.4). This framework thus provides a CAPM-equivalent decomposition of average total variance into market variance and average idiosyncratic variance, with the considerable advantage that it bypasses the need to estimate possibly time-varying *betas* for each asset.

The same scheme continues to apply if we introduce further levels of decomposition. CLMX (2001), decompose returns  $r_{ij,t}$  on firm  $i$  that belongs to industry  $j$  taken from the market portfolio  $m$  into industry and the firm-level components:

$$\begin{aligned}
 r_{ij,t} &= \beta_{ij,j} \varepsilon_{j,t} + \beta_{ij,m} r_{m,t} + e_{ij,t} \\
 &= r_{j,t} + \eta_{ij,t} \\
 &= r_{m,t} + \eta_{j,t} + \eta_{ij,t}
 \end{aligned} \tag{3.6}$$

Where<sup>26</sup>,

$$\begin{aligned}
 r_{j,t} &= \beta_{j,m} r_{m,t} + \varepsilon_{j,t} \\
 &= r_{m,t} + \eta_{j,t}
 \end{aligned}$$

Here,  $\varepsilon_{j,t}$  is the residual<sup>27</sup> of the regression (with no intercept) of industry  $j$  on the market excess-return and thus it is the component of the former orthogonal to the latter. Furthermore,  $\beta_{ij,j}$ ,  $\beta_{ij,m}$  and  $\beta_{j,m}$  are regression coefficients,  $\eta_{ij,t}$  is the firm-level, industry-adjusted component of  $r_{ij,t}$  and  $\eta_{j,t}$  is the industry-level, market adjusted component<sup>28</sup> of  $r_{j,t}$ . The average total variance  $var_t^{ind}$  of the  $n$  industries that belong to the market portfolio  $m$  can be decomposed, using (3.4), into the portfolio variance  $Var(r_{m,t})$  and an average industry level component  $IND_t$ :

<sup>26</sup> Also, since  $\varepsilon_{j,t}$  and  $r_{m,t}$  are orthogonal (by construction),  $\beta_{ij,m} = \beta_{ij,t} \beta_{i,m}$ .

<sup>27</sup> Notice that this is different from  $\varepsilon_{i,t}$  in (3.1), as the latter is a regression of asset  $i$  (instead of industry  $j$ ) return on the market return.

<sup>28</sup> Notice that this is different from  $\eta_{i,t}$  in (3.1), as the latter is the market-adjusted component of asset  $i$  return (instead of industry  $j$  return).

$$\begin{aligned}
var_t^{ind} &= \sum_{j=1}^n w_{j,t} Var(r_{j,t}) \\
&= Var(r_{m,t}) + \sum_{j=1}^n w_{j,t} Var(\eta_{j,t}) \\
&= Var(r_{m,t}) + IND_t
\end{aligned} \tag{3.7}$$

Similarly, the average total variance of the  $k$  firms that belong to industry  $j$  can be decomposed, using (3.4), into the industry variance  $Var(r_{j,t})$  and an average firm level component  $FIRM_{j,t}$ :

$$\begin{aligned}
var_{j,t} &= \sum_{i=1}^k w_{ij,t} Var(r_{ij,t}) \\
&= Var(r_{j,t}) + \sum_{i=1}^k w_{ij,t} Var(\eta_{ij,t}) \\
&= Var(r_{j,t}) + FIRM_{j,t}
\end{aligned} \tag{3.8}$$

Using (3.8) and averaging across industries, the average total variance of the firms that belong to the portfolio  $m$  can be decomposed into the average total industry variance,  $var_t^{ind}$ , and an average firm-level component  $FIRM_t$ .

$$\begin{aligned}
var_t &= \sum_{j=1}^n w_{j,t} var_{j,t} \\
&= \sum_{j=1}^n w_{j,t} Var(r_{j,t}) + \sum_{j=1}^n \sum_{i=1}^k w_{j,t} w_{ij,t} Var(\eta_{ij,t}) \\
&= var_t^{ind} + FIRM_t
\end{aligned} \tag{3.9}$$

In turn, using (3.7), the average total industry variance can be decomposed into the portfolio variance and the average industry level component  $IND_t$  and thus (3.9) becomes:

$$var_t = Var(r_{m,t}) + IND_t + FIRM_t \tag{3.10}$$



The variance components thus constructed share the same properties of the corresponding variance components derived from equation (3.4). The idiosyncratic industry and firm-level variance components, in particular, can be interpreted as the average variance of the residuals from the multiple regressions of stock returns on market returns and on residuals from the regression of industry returns on market returns, i.e.  $IND_t \cong \sum_{j=1}^n w_{j,t} \varepsilon_{j,t}$  and  $FIRM_t \cong \sum_{j=1}^n \sum_{i=1}^k w_{ij,t} e_{ij,t}$ . The upshot of this variance decomposition is that it provides an intuitively appealing interpretation of the nature of average idiosyncratic risk. It can be seen as the volatility of a diversified portfolio of market and industry neutral relative-value trades.

### 3.4. Unconditional and Conditional Second Moments Estimates

There are two main approaches to the estimation of the second moments of the distribution of asset returns. The first approach, ascribed to Officer (1973) and Merton (1980), is based on unconditional estimation methodology. The second approach, initiated by the seminal paper of Engle (1982), uses a conditional estimation methodology to model conditional heteroschedastic patterns.

#### 3.4.1 Unconditional Estimates

The unconditional estimation approach treats the moments of a distribution as an observable variables and it therefore uses unconditional sampling methodologies to construct volatility and correlation estimates. Consider the following model of a multivariate stochastic process  $y_t$  with possibly time-varying moments:

$$y_t = \mu_T + u_t \tag{3.11}$$

$$u_t \sim N(0, H_T)$$

Where:

$$H_T \equiv D_T C_T D_T$$

$$\mu_T = E_{t \in T}(y_t)$$

$t = 1, 2, \dots, p$  sub-periods of  $T$

Here,  $\mu_T$  is the time-varying mean of  $y_T$ ,  $C_T$  is a symmetric correlation matrix with ones along the main diagonal,  $D_T$  is the diagonal matrix of standard deviations and  $u_t$  is the  $n \times 1$  vector of zero mean innovations, obtained by subtracting the means over the  $p$  sub-periods in period  $T$  from each one of the  $n$  elements of  $y_t$  and stacking them. Multivariate normality is not strictly necessary but I assume it for ease of exposition. I also assume that  $C_T$  is positive-definite and  $D_T$  is positive semi-definite, thus  $H_T$  is positive semi-definite too.

In the unconditional estimation methodology, moments are treated as observable random variables and polynomials of returns are seen as their realizations. If the distribution of the moments is assumed to be stable, the corresponding polynomials can be interpreted as stationary processes. Thus, by some central limit theorem, their mean exists and it can be consistently estimated using the sample average of their realizations. Asymptotically, it can be estimated to any desired degree of accuracy by increasing the sample size.

In particular, within each period  $T$ , squared returns innovations,  $u_{i,t}^2$ , and innovations cross-products,  $u_{i,t}u_{j,t}$ , are seen as realizations drawn from given underlying distributions of volatilities and correlations. A central limit theorem therefore applies and the elements of  $D_T$  and  $C_T$  can be estimated with greater and greater accuracy by drawing increasingly large samples of squared returns and returns cross-products, respectively. In practice, the sample size used for the estimation of volatilities and correlations over a particular period can be increased by sampling squared returns and returns cross-products at arbitrarily high frequencies (by increasing  $p$ ).

Once estimates of the elements of the  $D_T$  and  $C_T$  matrices have been computed, the variance-covariance matrix is obtained by simply feeding them into  $H_T$ . This procedure yields volatility, correlations, variances and covariances estimates for each period  $T$ . This is sometimes called the ‘rolling’ volatility and correlation



estimator (see, for example, Engle and Mezrich (1996)). These estimates are then typically grouped into sequences of non-overlapping, consecutive periods  $T$  to form estimated second moments time series. Formally, dividing a period  $T$  into  $p$  sub-periods the variance of the zero-mean returns innovations  $u_t$  is:

$$Var(u_t)_T = \frac{\sum_{t=1}^p u_t^2}{p} \quad (3.12)$$

And for correlations between returns innovations  $u_{i,t}$  and  $u_{j,t}$ :

$$c_{i,j,T} = \frac{\sum_{t=1}^p u_{i,t} u_{j,t}}{\sqrt{Var(u_{i,t})_T Var(u_{j,t})_T}} \quad (3.13)$$

Since this approach treats the volatility and correlation processes as observable variables, inferences about their behavior over time can be made by studying their distribution. This approach has been recently applied by Andersen, Bollerslev, Diebold and Ebens (2000) to the study of the behaviour over time of volatility and correlation using high-frequency intraday return data. In this approach, it is also legitimate to study the distribution of the volatility and correlation processes conditional on other variables. This amounts to drawing inferences about their behavior by fitting regression models to the constructed unconditional estimates. CLMX (2001) construct their market and idiosyncratic volatility series using an unconditional estimation approach. In particular, they compute market volatility and the individual stocks firm-level variances needed to construct average industry and firm-level variance as the sample sum of, respectively, daily squared returns and market adjusted returns over non-overlapping monthly periods. They then study the behaviour over time of the constructed series by testing for the presence of unit roots, by fitting univariate models with a deterministic time trend and by including these variables in multivariate VAR systems.

This strategy can yield useful insights, and it has the additional advantage of simplicity. But it has three shortcomings. *First*, even though the sums of squares and cross-products are consistent estimators of the second moments of the return distributions at each point in time, there might be considerable bias in finite samples (especially in relatively small samples) since they are *ad hoc* representations of the volatility and correlation processes. In particular, all observations within the last  $p$  periods are given equal weight and observations older than  $p$  periods are given zero weight. This gives rise to ‘shadows’ (spikes in the volatility and correlation estimates up to  $p$  periods after major events) and there is no theoretical guidance as to how to choose  $p$ . The theory only says that  $p$  should be chosen as high as possible to increase accuracy (to reduce the variance of the estimate according to the relevant central limit theorem), thus neglecting the importance of temporal aggregation issues. The latter arise from the circumstance that the parameters and their values that best describe the process might change depending on the frequency. An example is the different volatility implied at different horizons by the returns auto-correlation structure, e.g. if auto-correlation is negative (positive), low frequency volatility is lower (higher) than high frequency volatility. *Second*, the aggregation of daily data into lower frequency monthly data leads to a potential small sample problem. *Third*, there is no guarantee that the estimated variance-covariance matrix turns out positive-definite.

Care should be taken when using the estimated second moment series as explanatory variables to avoid a possible ‘generated regressors’ problem. This is a special “error in variables” problem discussed by Pagan (1984) and Pagan and Ullah (1988). In employing volatility and correlation estimates as explanatory regression variables, this problem arises if there are reasons to believe that their variance is large. This would be the case if the sample size  $p$  is small relative to the variance of the underlying sampled variable<sup>29</sup>. Other important issues arise when employing

---

<sup>29</sup> The variance of the estimated volatility and correlation series is, by the relevant central limit theorem, equal to the variance of the corresponding underlying variable divided by the sample size  $p$ . For example, the variance of the estimated volatility of asset  $i$  is  $Var(\sigma_{iT})/p$ , where  $Var(\sigma_{iT})$  is the ‘true’ variance of the variance process. Since  $Var(\sigma_{iT})$  is unlikely to be known *a priori*, its sample



estimated second moment series as explanatory variables of returns, as in tests of Merton's (1973, 1980) ICAPM that regress aggregate returns on lagged market variance. As discussed by Pagan (1984), the use of lagged realized variance as a regressor in place of the unobservable rational expectation of the variance leads to consistent estimates of the coefficient on the latter, under the assumption that unexpected returns (the residuals of the regression model) are independently distributed. In this case, consistent standard errors should be computed using standard heteroskedasticity adjusted variance covariance matrix estimators, such as White's (1980). If the residuals are serially correlated, standard errors should be corrected further, e.g. using the Newey-West's (1987) variance-covariance matrix estimator. Merton (1980) also shows that constructing the second moment series using overlapping returns generates contemporaneous correlation between the regressor and the regression residual. In the presence of serial correlation in the regression residual, this leads to inconsistent regression estimates.

### *3.4.2 Conditional Estimates*

The second approach to the estimation of the moments of the return distribution, initiated by the pioneering work of Engle (1982), treats them as latent variables to be estimated conditional on the available information set. In the discussion that follows, I will focus on fully parametric specifications, where the statistical model of the variables of interest is completely known up to some parameters. They include Autoregressive Conditional Heteroskedasticity (ARCH) and Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models. These models provide a specification for the dynamics of the conditional second moments of the noise component of a random variable time-series representation. They are the extension to second moments of essentially analogous models used for the mean, such as the Autoregressive (AR) and Autoregressive Moving Average (ARMA) models (Box and Jenkins (1970)). For the discussion of conditional second moment

---

estimate could be used instead. Assuming that the estimated underlying volatility process is stationary, sample estimates could be formed by computing the sample variance of the volatility estimates.

models and estimation methodologies that follows, it is preliminarily useful to specify the conditional multivariate distribution of the vector of random variables  $y_t$ :

$$\begin{aligned} y_t &= E_{t-1}(y_t) + u_t \\ u_t \mid \mathfrak{I}_{t-1} &\sim \Phi(0, H_t) \end{aligned} \quad (3.14)$$

Where,

$$H_t \equiv D_t C_t D_t$$

Here, symbols retain their prior meanings but the time subscripts change. Also, importantly, I work with conditional instead of unconditional moments. Briefly,  $u_t$  is an  $n \times 1$  vector of zero mean innovations conditional on the information set available at time  $t-1$  ( $\mathfrak{I}_{t-1}$ ), obtained by subtracting the conditional means from each of the  $n$  elements of  $y_t$  and stacking them. They follow a  $\Phi$  distribution, not necessarily normal, with centred second moment matrix  $H_t$ . Also,  $C_t$  is the conditional correlation matrix and  $D_t$  is the diagonal matrix of conditional standard deviations. Again, both  $D_t$  and  $C_t$  and, as a consequence,  $H_t$  are assumed to be positive definite. To further specify the conditional statistical model of  $u_t$ , I parameterize its moments as functions of the finite dimensional vector  $\psi$ . Therefore we have  $u_t = u_t(\psi)$ . As before, the  $y_t$  vector could represent the returns on the assets that constitute a portfolio.

ARCH models were introduced by Engle (1982) and are now commonly used to describe and forecast changes in the volatility of financial time series. They are useful when we have reason to believe that the variance of the error term varies over time as a function of how large the errors were in the past. Formally, the  $n \times 1$  vector stochastic process  $u_t$  follows a multivariate ARCH process if the  $n \times n$  conditional variance-covariance matrix depends non-trivially on the past of the process:

$$H_t = E_{t-1}(u_t u_t' \mid \mathfrak{I}_{t-1}; \psi) = E_{t-1}(u_t u_t' \mid u_{t-1}, u_{t-2}, u_{t-3}, \dots, u_1; \psi) \quad (3.15)$$



Notice that the  $u_t$  innovations are clearly non independently distributed once we allow conditional second moments to depend on past realizations of the process. Now define the standardized innovations as  $z_t = H_t^{-1/2}u_t$  (one of the good reasons to require  $H_t$  to be positive-definite is to be able to invert it). These will have zero mean and time-invariant unit variance and zero covariance, i.e. their variance-covariance matrix is the identity matrix  $I_n$ . This observation is central to most of the inference procedures based on ARCH-type models. Specification checks, in particular, are aimed at ascertaining whether the vector  $z_t$ , computed using the estimated variance-covariance matrix, displays any residual time pattern in its second moments or any correlation between its elements.

Allowing for time-variation in conditional second moments implies excess-kurtosis in the unconditional distribution. In other words, it allows modelling the phenomenon of fat tails observed in the distribution of many financial and economic variables. To see this, consider for expositional simplicity the univariate case. If the distribution of  $z_t$  is assumed to be time-invariant with finite fourth moment, the unconditional fourth moment of  $u_t$  can be written as the product between the unconditional fourth moment of  $z_t$  and the unconditional second moment of the conditional volatility rate:

$$E(u_t^4) = E(E(z_t^4 \sigma_t^4 | \mathfrak{F}_{t-1})) = E(z_t^4)E(\sigma_t^4) \quad (3.16)$$

Then, since the square is a convex function of a variable, it follows by Jensen's inequality that the mean of the square of  $\sigma_t^2$  is greater than the square of its mean. Therefore:

$$E(z_t^4)E(\sigma_t^4) \geq E(z_t^4)E(\sigma_t^2)^2 \quad (3.17)$$

Finally, since by definition  $E(u_t^2) = E(\sigma_t^2)$ , we can substitute and write:

$$E(u_t^4) = E(z_t^4)E(\sigma_t^4) \geq E(z_t^4)E(u_t^2)^2 \quad (3.18)$$

Here the equality holds true only for constant conditional variance, see for example Bollerslev, Engle and Nelson (1994). This means that time-variant conditional variance implies that the unconditional fourth moment of  $u_t$  is larger than the unconditional fourth moment of  $z_t$  scaled up by the appropriate measurement unit, i.e.  $E(u_t^2)^2$ , and thus that the innovations have an unconditional distribution characterized by excess-kurtosis even if the standardised innovations  $z_t^4$  are not leptokurtic. In particular, even if the conditional distribution of the innovations is Gaussian, the unconditional distribution is non-Gaussian, with excess-kurtosis due to the mixture of Gaussian densities with different volatilities.

The second moment parameters are usually estimated by maximum likelihood (ML). This involves finding the values  $\hat{\psi}$  that maximize the likelihood function of the sample of  $T$  realizations  $\{y_{T-1}, y_{T-2}, y_{T-3}, \dots, y_1\}$  of the multivariate process followed by  $y_t$  under the null that  $u_t$  follows the hypothesized ARCH-class process. In practice, the strictly monotone logarithmic transformation of the likelihood function is usually maximized. The log-likelihood of the  $t$ th observation is<sup>30</sup>:

$$l_t(y_t; \psi) = \ln f[z_t(\psi)] - 0.5 \ln(|H_t(\psi)|) \quad (3.19)$$

Defining  $L_T$  as the log-likelihood of the full sample of  $T$  observations, the estimation problem is then:

$$\text{Max}_{\psi} L_T(y_T, y_{T-1}, y_{T-2}, \dots, y_1; \psi) \quad (3.20)$$

In the normally distributed case or under the auxiliary assumption that  $z_t$  is *i.i.d.*, the sample log-likelihood simplifies to:

$$L_T(y_T, y_{T-1}, y_{T-2}, \dots, y_1; \psi) = \sum_{t=1}^T l_t(y_t; \psi) \quad (3.21)$$

---

<sup>30</sup> The second term on the right is a Jacobian that arises in the transformation from the standardized innovations to the observable  $y_t$ .



The normality assumption is often made for the conditional distribution of  $y_t$ ,  $u_t$  and, therefore, the standardized innovations  $z_t$ . Even though this assumption might appear implausible in many empirical applications, e.g. the estimation of the dependence structure of financial return data with non-normal distributional features such as thick tails and high skewness and kurtosis, it might be alternatively justified on quasi-maximum likelihood (QML) grounds, following Bollerslev and Wooldridge (1992). If the conditional mean and variance equation are specified correctly, the QML estimator obtained by maximizing a conditional normal likelihood function is Fisher-consistent, no matter what the true distribution of the population from which the observations are drawn might be. Under appropriate regularity conditions, this is sufficient to establish consistency and asymptotic normality of the parameters estimates,  $\hat{\psi}$  (see, for a discussion, Bollerslev, Engle and Nelson (1994)). Moreover, while using distributions other than the normal might yield more efficient estimates if the distributional assumption is correct and the conditional model is correctly specified, it might end up sacrificing consistency when this is not the case, see Fiorentini, Sentana and Calzolari (2002) for a discussion. When non-normal distributions are assumed, moreover, theoretical results about the asymptotic distribution of the estimators might not be readily available. The log-likelihood function under the multivariate normality assumption is:

$$\begin{aligned}
 L_T(\psi) &= -\frac{1}{2} \sum_{t=1}^T [y_t - E_t(y_t)]' H_t^{-1}(\psi) [y_t - E_t(y_t)] - \frac{Tn}{2} \ln(2\pi) - \frac{Tn}{2} \ln |H_t(\psi)| \\
 &= -\frac{1}{2} \sum_{t=1}^T u_t' H_t^{-1}(\psi) u_t - \frac{Tn}{2} \ln(2\pi) - \frac{Tn}{2} \ln |H_t(\psi)| \\
 &= -\frac{1}{2} \sum_{t=1}^T z_t(\psi)' I_t^{-1} z_t(\psi) - \frac{Tn}{2} \ln(2\pi) \tag{3.22}
 \end{aligned}$$

I will now describe a number of conditional second moment models belonging to the ARCH class. These are specifications of (3.13) that impose particular restrictions on the temporal dependency in the conditional mean and variance equations. I will

mainly focus on the latter. I start from the univariate case, and then I generalize the discussion to multivariate models.

### *Univariate Models*

Suppose that returns are described by the following regression model:

$$R_t = \gamma_0 + \gamma_1 x_{1t} + \dots + \gamma_n x_{nt} + u_t \quad (3.23)$$

Here,  $E_{t-1}(u_t) = 0$  and  $E_{t-1}(u_t^2) = \sigma_t^2$ . Suppose further that we can then write a second equation relating the variance of the error term to the amount of volatility observed in past periods:

$$\sigma_t^2 = a_0 + a_1 u_{t-1}^2 + \dots + a_p u_{t-p}^2 \quad (3.24)$$

Here, there are two components of the conditional variance of  $u_t$ , a constant term and previous periods news about volatility (squared residuals). In other words, the conditional variance is parameterized as an autoregressive distributed lag of  $p$  squared innovations. The residual is thus heteroskedastic, conditional on  $u_t, u_{t-1}, \dots, u_{t-p}$ . This model is known in the literature as ARCH( $p$ ), where  $p$  is the number of autoregressive terms in the conditional variance equation. This is the ARCH specification originally introduced by Engle (1982). The  $p+1$  parameters  $a_0, a_1, \dots, a_p$  of the conditional variance equation must be estimated along with the parameters  $\gamma_0, \gamma_1, \dots, \gamma_n$  of the conditional mean equation. The parameter estimates are the values  $\hat{\psi}$  that maximize the likelihood function of the sample of  $T$  realizations  $\{R_{T-1}, R_{T-2}, R_{T-3}, \dots, R_1\}$ .

Since (3.24) can be seen as a distributed lag model for  $\sigma_t^2$ , we can replace many lagged values of  $u_t$  (the conditional residuals) by only one or two lagged values of  $\sigma_t^2$  (the conditional variance). In general, we could have any number of ARCH terms



and any number of lagged values of the conditional variance in the specification of a GARCH model (Bollerslev (1986)), although parsimony is generally recommended. In the conditional variance equation we could also include explanatory variables from the conditional mean equation. This is shown in the following general representation of a GARCH( $p, q$ ) model:

$$R_t = \gamma_0 + \gamma_1 x_{1t} + \dots + \gamma_n x_{nt} + e_t \quad (3.25)$$

$$\sigma_t^2 = a_0 + a_1 u_{t-1}^2 + \dots + a_p u_{t-p}^2 + b_1 \sigma_{t-1}^2 + \dots + b_q \sigma_{t-q}^2 + \varphi_i x_{it} \quad (3.26)$$

Again, all the parameters in both the conditional mean equation and the conditional variance equation must be estimated simultaneously by ML or QML. In order to rule out negative variances, both sides of all the conditional variance equations should be non negative. In the GARCH( $p, q$ ) model, (3.26) with  $\varphi_i = 0$ , the unconditional variance estimate is the following:

$$\sigma^2 = \frac{a_0}{1 - \sum_{i=1}^p a_i - \sum_{i=1}^q b_i} \quad (3.27)$$

This is the base-line level to which variance mean reverts in the long-run if it follows a stationary process. A necessary condition for variance to be stationary, therefore, is that  $\sum_{i=1}^p a_i + \sum_{i=1}^q b_i \neq 1$ . By a similar reasoning, the latter expression represents the persistence of the conditional variance estimates. While a non-explosive variance process requires  $\sum_{i=1}^p a_i + \sum_{i=1}^q b_i \leq 1$ , the closer this expression is to unity, the more persistent are the effects of innovations on subsequent values of the series.

Notice that the dynamic behaviour of the conditional second moments, under any specification that belongs to the ARCH class, implies a temporal aggregation problem. If a GARCH model is correctly specified for one frequency of data, then it

is in general misspecified for data with any other time scale (Engle and Patton (2001)). In other words, the circumstance that the conditional variance of a process is generated by an ARCH-type model at a particular frequency does not imply that the data at different frequencies be generated by the same model. The consequence is that parameters estimates using data with different frequencies might imply different rates of decay of volatility innovations (see for example Kearney and Patton (2000)).

Constraining the  $\sum_{i=1}^p a_i + \sum_{i=1}^q b_i$  expression to sum to unity yields the integrated GARCH( $p,q$ ), or IGARCH( $p,q$ ), specification. This constraint forces the conditional variance to act, in some respects, like a unit process. In particular, the one-step ahead forecast of the IGARCH(1,1) variance depends only on the constant term and the past value of conditional variance but, unlike a true unit root process, each realization of the series is a geometrically decaying function of the current and past squared innovations. In this model unconditional variance is undefined. See Nelson (1990) for a discussion. The IGARCH(1,1) process, with  $a_0$  in (3.26) constrained to be zero, yields the Exponential Moving Average (EWMA) representation of the variance process. Because of its computational simplicity, coupled with its ability to capture the high persistence of stock returns volatility, this representation is very popular in the financial industry (e.g. JP Morgan's RiskMetrics™ approach).

Since understanding and predicting the temporal dependence in the second order moments of asset returns is important for many issues in financial econometrics, a bewildering family of univariate GARCH models has proliferated. These models are extensions of the basic GARCH( $p,q$ ) model in (3.25) and (3.26). These extensions address a number of important stylized empirical regularities about volatilities such as leverage and asymmetric volatility effects, conditionally fat tails and the possibility of regime switches.



Typically, as summarized by the leverage and feed-back volatility effects highlighted by Black (1976), and Christie (1982), Schwert (1989), Campbell and Hentschel (1992) and discussed in Section 3.2.3, negative innovations lead to larger volatility increases than positive innovations. In models that allow for this type of asymmetry, conditional volatility depends not only on the size, but also on the sign of the innovations. This can be achieved by either adding a new predetermined economic variable, e.g. firms' financial leverage, to the variance equation or by modifying the autoregressive component of the process. A simple example of the latter approach is the model estimated by Glosten, Jagannathan and Runkle (1993). They add the product between the lagged squared innovation and an indicator variable to a standard GARCH(1,1) specification. The indicator variable takes the value 1 or 0 depending on whether past returns are, respectively, negative or positive. Their variance equation takes the following form:

$$\sigma_t^2 = a_0 + a_1 u_{t-1}^2 + a_2 (u_{t-1}^2 I_{t-1}) + b_1 \sigma_{t-1}^2 \quad (3.28)$$

Here,  $I_{t-1} = 1$  if  $u_{t-1} < 0$ ,  $I_{t-1} = 0$  otherwise. If the innovation  $u_{t-1}$  is negative the effect on  $\sigma_t^2$  is  $(a_1 + a_2)u_{t-1}^2$ , otherwise the effect is  $a_1 u_{t-1}^2$  as in the symmetric GARCH(1,1) case. Henceforth, I will denote this specification as GJR-GARCH. A more complex asymmetric GARCH process is the exponential GARCH (E-GARCH) model proposed by Nelson (1991). The E-GARCH variance process is formulated in terms of the strictly non decreasing logarithmic transformation of the conditional variance. This specification has the advantage that the volatility estimates are never negative, no matter the value of the parameters estimates. The conditional log-variance equation is:

$$\ln(\sigma_t^2) = a_0 + a_1 \left| \frac{u_{t-1}}{\sigma_{t-1}} \right| + a_2 \frac{u_{t-1}}{\sigma_{t-1}} + b_1 \ln(\sigma_{t-1}^2) \quad (3.29)$$

If the standardized innovation  $\frac{u_{t-1}}{\sigma_{t-1}}$  is negative, the effect on the logarithm of the conditional variance is  $a_1 - a_2$ , otherwise the effect is  $a_1 + a_2$ . Other specifications are possible. Engle and Ng (1993) review a number of alternative specifications and introduce the news impact curve. This curve is derived conditioning on the information available at  $t - 2$  in order to consider the effect of the shock  $u_{t-1}$  on the conditional variance  $\sigma_t^2$  in isolation. Different ARCH and GARCH models can thus be compared by asking how the conditional variance is affected by the latest information, the ‘news’. This is particularly useful in order to visualize how the various specifications model asymmetries in the variance process. For example, the news impact curve of the GARCH(1,1) model has the form  $\sigma_t^2 = A + a_1 u_{t-1}$ , where  $A = a_0 + b_1 \sigma^2$  (recall that  $\sigma^2$  is the unconditional variance). This curve is symmetric with respect to  $u_{t-1} = 0$ . Other GARCH models, such as the GJR-GARCH and the EGARCH, have asymmetric news impact curves. See Engle and Ng (1993) and Ding, Granger and Engle (1993) for examples and a discussion.

One further extension of the basic GARCH( $p,q$ ) model is represented by GARCH in mean (GARCH-M) specifications originally proposed by Engle, Lilien and Robins (1987). Consistent with the predictions of popular asset pricing models such as Merton’s (1973) ICAPM, GARCH-M models allow for the conditional variance in (3.26) to enter the mean of the process in (3.25). Because they explicitly model the conditional risk-return relationship, GARCH-M specifications can be used to account for asymmetries in the distributions of financial returns introduced by the volatility feed-back effect (Campbell and Hentschel (1992)). For example, Chou (1988) estimates a GARCH-M model of returns at various frequencies (from weekly to annual) over the period 1962-1985. The estimated excess-return mean equation is (3.23) with  $x_{1t} = \sigma_t$  and  $\gamma_2 = \gamma_3 \dots = \gamma_n = 0$ . The estimated variance equation is (3.26) with  $p = q = 1$ . He finds plausible point estimates of the ‘relative risk aversion’ parameter  $\gamma_1$  but the latter is only marginally statistically significant for most returns horizon. Attanasio and Wadhvani (1990) use a GARCH-M specification to model the time-varying US stock-market excess return over the period 1953-1988. They



then test whether, against the EMH, either the dividend-price ratio (or dividend yield), a short term interest rate or both predict unexpected stock market excess returns. Their estimated market price of risk is both uncomfortably large from a theoretical point of view (under power utility, it implies an implausibly high risk aversion) and only marginally significant. Moreover, the short rate does appear to predict conditionally unexpected excess returns. This, however, might be rationalized under Merton's (1973) ICAPM by interpreting the short rate as a variable that proxies for the state of the investment opportunity set (Scruggs (1998)).

Another possibility is to estimate the GARCH model assuming standardized Student-t distributed conditional innovations as in Bollerslev (1987). As the degrees of freedom go to infinity, this distribution approaches the standard normal. While the ability to capture the second moment dynamics of unconditional distributions with fatter tails than under the normality assumption is typical of all the models that belong to the ARCH family, Student-t distributed conditional innovations allow for fatter tails of the conditional distribution. More specifically, while the standard GARCH model in (3.26) with normally distributed conditional innovations displays by (3.16) unconditional excess-kurtosis, it also displays excess-kurtosis in the conditional distribution of the innovations if these are Student-t distributed. Obviously, the latter specification allows for larger excess-kurtosis in the unconditional distribution. This specification is particularly useful for modelling returns on assets that experience large gains or large losses relatively often (more often than under a normal distribution), such as some stocks traded in emerging markets, certain emerging stock markets indices and, among non-equity variables, foreign exchange rate returns.

Although ARCH and GARCH type models estimated assuming a t-distribution for the conditional innovations can account for fat tails, such specifications will typically predict too much volatility persistence in the presence of a regime switch. Switching ARCH/GARCH models were first proposed by Hamilton and Susmel (1994) to let persistence in volatility depend not only on the size of past residuals



and conditional variance but also on the number and characteristics of regimes or states. The transition between different volatility states is governed by a Markov process. This mechanism is designed to capture two stylized facts that conventional generalized autoregressive conditional heteroskedasticity models find hard to reconcile, namely that conditional volatility can increase substantially in a short amount of time at the onset of a turbulent period while its rate of mean reversion appears to vary positively and nonlinearly with its level. In other words, stock-market volatility does not remain persistently two to three times above its normal level the same way it persists at a level that is just 30-40 percent above normal. Dueker (1997) allows the student-t degrees-of-freedom parameter to switch according to a Markov process such that the conditional variance and kurtosis are subject to discrete shifts and he finds that the half life of the most leptokurtic state is relatively short. One implication is that expected market volatility reverts to near-normal levels fairly quickly following a spike. Edwards and Susmel (2000) estimated both GARCH(1,1) and SWARCH (1,1) models<sup>31</sup> of weekly interest rates for Argentina, Brazil, Chile, Hong Kong and Mexico from 1994 to 1999 allowing for three different volatility states (low-medium-high). The estimated parameters imply an explosive volatility process in the case of the GARCH(1,1) models but not in the case of the SWARCH(1,1) models.

One way to test for ARCH-type effects is to estimate an ARCH-type model of the second moment process and then perform significance t-tests or F-tests (typically for the null hypothesis that either some or all the parameters of the model are not significantly different from zero). This amounts to checking on the fit of the estimated model. An estimated GARCH model, however, should account for all the dynamic aspects of the mean and variance process. The estimated conditional residuals should be uncorrelated and should not display conditional heteroskedasticity in their second moments. Thus, the specification of the model can be tested by first forming the standardised conditional innovations as  $z_t = \frac{u_t}{\sigma_t}$ . If

---

<sup>31</sup> They also estimate bivariate versions, for two countries at a time, of these models.



there is any residual serial correlation in  $z_t$  the model for the mean is not correctly specified. On the other hand, the presence of patterns in the square of  $z_t$  suggests the presence of residual heteroskedasticity and, hence, that the model for the variance is not correctly specified. One way to test for patterns of a linear nature in the volatility of the conditional standardised residuals is to use the Ljung-Box Q statistic (see Enders (2004)) of their square. To check on the presence of residual patterns of a possibly non linear nature one could use Ramsey's RESET test or Brock, Dechert, Scheinkman (1996) BDS test on the squares of the conditional standardized residuals.

One way to test for leverage effects or for asymmetric reaction of volatility to news is to estimate a GJR-GARCH or an EGARCH model and then perform a t-test for the null hypothesis that  $a_2 = 0$  in (3.28) and (3.29). However, there is a specific diagnostic test to determine whether there is a residual leverage effect not captured by the estimated conditional variance model. This test requires, after estimating the selected ARCH or GARCH specification, to construct the standardised conditional innovations  $z_t$ . Then, to test for residual leverage effects, the following regression should be preliminarily estimated:

$$z_t^2 = \delta_0 + \delta_1 z_{t-1} + \delta_2 z_{t-2} + \dots + \delta_k z_{t-k} + \varepsilon_t \quad (3.30)$$

Here,  $\varepsilon_t$  is a i.i.d. regression error term. If there are no residual leverage effects, the squared standardized errors should be uncorrelated with lags of the standardized errors. Thus the null hypothesis that there are no residual leverage effects is that  $\delta_0 = \delta_1 = \dots = \delta_k = 0$ . Under the appropriate regularity conditions, this can be tested using a standard F-test. This is sometimes known as a 'size bias' test. Engle and Ng (1993) proposed another procedure that can be used as a misspecification test for GARCH models under the null of no residual asymmetry in the response of second moments to innovations of different signs. They suggest to test whether the estimated squared residuals can be predicted using the indicator variable  $I_{t-1}$  that

takes value 1 if the lagged innovation is negative and zero otherwise. This is known as the ‘sign bias’ test. This is based on the estimation of the following regression:

$$z_t^2 = \delta_0 + \delta_1 I_{t-1} + \varepsilon_t \quad (3.31)$$

Under the null of the test, there are no residual asymmetric effects in the errors of the model. If a  $t$ -test indicates that  $\delta_1$  is significantly different from zero, the sign of the current period shock is helpful in predicting next period conditional variance and the null is rejected. This test can be extended to determine whether the effects of positive and negative innovations also depend on their size:

$$z_t^2 = \delta_0 + \delta_1 I_{t-1} z_{t-1} + \varepsilon_t \quad (3.32)$$

This is known as the ‘negative size bias’ test. Under the null of the test, the size of the innovation does not influence whether the response of volatility is asymmetric or not. These tests can be generalized further. In particular, Engle and Ng (1993) propose a joint test for sign and size bias. This requires that the following regression be estimated:

$$z_t^2 = \delta_0 + \delta_1 I_{t-1} + \delta_2 I_{t-1} z_{t-1} + \delta_3 (1 - I_{t-1}) z_{t-1} + \varepsilon_t \quad (3.33)$$

In this equation,  $(1 - I_t)$  takes by construction value 1 when the lagged innovation is positive and zero otherwise. Significance of  $\delta_1$  suggests the presence of sign bias, the significance of  $\delta_2$  and  $\delta_3$  suggest the presence of negative and positive size bias, respectively.

### *Multivariate Models*

Financial volatilities are not only time-varying, but they also move together across assets and markets (Bollerslev, Engle and Nelson (1994)). Given a set of assets (e.g.



a portfolio), estimating their conditional second moment process using univariate ARCH type models imposes the restrictions that the individual asset variances do not depend on anything but their own lagged squared return innovations and their own lag. In particular, individual asset variances are not allowed to depend on lagged squared innovations of the return on other assets, on the cross-products of the latter, or on other variances and other correlations. Recognizing the possibility of co-movements in a multivariate modelling framework leads to obvious gains in efficiency and, therefore, it leads to more reliable empirical representations than working with separate univariate models. If time variation in co-movements displays persistence and clustering like the time variation in volatilities, multivariate generalizations of the ARCH type class of models are obvious candidates to model them (Bollerslev, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994)).

The development of multivariate ARCH type models opens the door to better decision making tools in various financial applications such as asset pricing models, portfolio selection, option pricing and Value-at-Risk estimation. As noticed by Engle and Kroner (1995), the extension of univariate conditional second moment models to the multivariate case is conceptually analogous to the extension of ARMA models of the mean to vector ARMA ones. Like univariate specifications, they can be estimated with either normal (again, in a ML or QML setting) or non-normal distributions. The rationale for using a parametric leptokurtic distribution such as the Student-t distribution in a multivariate context is that one may be interested in the probability of joint occurrence of several extreme events, which is regularly underestimated by the multivariate normal distribution, especially in larger dimensions (Fiorentini, Sentana and Calzolari (2002)).

The estimation of multivariate second moments models, however, presents special computational difficulties. If more than just a few variables are jointly modelled, the number of their inter-relationships and, therefore, of free parameters to be simultaneously estimated can be very large. For example, a full unrestricted  $n$ -variable generalization of a GARCH (1,1) model, the so called *Vec* representation

(see for a review, Bollerslev, Engle and Nelson (1994)) and Engle and Kroner (1995)), requires the estimation of  $n^2$  equations and thus of  $O(n^4)$  parameters. Even though many of these parameters appear in redundant equations (for example, by symmetry of the variance covariance matrix, there are two identical covariance equations for each pair of variables), there are still  $n(n + 1)/2$  equations to be simultaneously estimated once the system is purged of the redundant ones. Formally, the  $Vec(p,q)$  representation, purged of the redundant equations, is the following:

$$h_t = \bar{h} + \sum_{j=1}^p A_j \eta_{t-j} + \sum_{j=1}^q B_j h_{t-j} \quad (3.34)$$

Where:

$$h_t = vech(H_t)$$

$$\eta_t = vech(u_t u_t')$$

Here,  $vech()$  is the operator that stacks the elements in the lower triangular portion of an  $n \times n$  matrix into an  $n(n+1)/2$  vector,  $A_j$  and  $B_j$  are conformable parameter matrices and  $\bar{h}$  is a  $n(n+1)/2$  parameter vector. The  $Vec$  representation is considerably general. All variances and covariances are a function of their own lags, of their own lagged squared innovation and innovation cross-products, of the lags on the other variances and covariances and of the other lagged squared innovation and innovation cross-products. However, this representation is also computationally very demanding because free parameters to be simultaneously estimated are attached to each right-hand side variable. To hasten the estimation procedure by reducing the number of parameters to be estimated, the elements of  $\bar{h}$  could be set equal to the corresponding elements of the sample estimate of the unconditional variance-covariance matrix. This technique is called ‘variance targeting’ and its usefulness is directly related to the assumption that the conditional second moments are a stationary process that mean-revert to a long-run mean given by the elements of  $\bar{h}$ . The unconditional variance-covariance estimate of  $H$  is the following:



$$\begin{aligned} \text{Vec}(\bar{H}) &= \left( I_{n(n+1)/2} - \sum_{j=1}^p A_j - \sum_{j=1}^q B_j \right)^{-1} \bar{h} \\ &= \left( I_{n(n+1)/2} - \Pi \right)^{-1} \bar{h} \end{aligned} \quad (3.35)$$

This is the base-line level to which the conditional variance-covariance matrix mean-reverts in the long-run if its elements follow a stationary process. A necessary condition for this is that the term in brackets has non-zero determinant and, therefore, that it is invertible. The necessary and sufficient condition for  $H_t$  to be stationary (covariance-stationary) is that the eigenvalues of  $\Pi$  are all less than one in modulus.

An alternative specification, the BEKK model introduced by Engle and Kroner (1995), imposes positive-definiteness of the estimated conditional variance-covariance matrices. Formally, the BEKK( $p,q,K$ ) model is defined as follows:

$$H_t = G'G + \sum_{k=1}^K \sum_{j=1}^p A'_{jk} u_{t-j} u'_{t-j} A_{jk} + \sum_{k=1}^K \sum_{j=1}^q B'_{jk} H_{t-j} B_{jk} \quad (3.36)$$

Here, all the matrices are  $n \times n$  and  $G$  is upper triangular. As proven by Engle and Kroner (1995), the BEKK( $p,q,K$ ) model is a special case of the *Vec* representation. They provide a number of propositions that specify the conditions under which the BEKK( $p,q,K$ ) model achieves full generality and it is therefore equivalent to a *Vec* representation. The vectorized expressions for the unconditional variance-covariance estimate  $\bar{H}$  are the following:

$$\text{Vec}(\bar{H}) = \left[ I_{n^2} - \sum_{k=1}^K \sum_{j=1}^p (A_{jk} \otimes A_{jk})' - \sum_{k=1}^K \sum_{j=1}^q (B_{jk} \otimes B_{jk})' \right]^{-1} \text{vec}(G'G) \quad (3.37)$$

$$\text{Vech}(\bar{H}) = \left( I_{n(n+1)/2} - \Pi \right)^{-1} \text{vech}(G'G) \quad (3.38)$$

As usual, this is the base-line level to which the conditional variance-covariance matrix mean-reverts in the long-run if it follows a stationary process. The condition for it to be stationary (covariance-stationary) is, as before, that the eigenvalues of  $\Pi$  are all less than one in modulus. Adopting a “variance targeting” approach, the  $G'G$  matrix could be set equal to the positive definite sample estimate of the unconditional variance-covariance matrix times the quantity  $(I_{n(n+1)/2} - \Pi)$ .

A somewhat more parsimonious but less general representation that requires the estimation of  $O(n^2)$  parameters is the BEKK( $p,q,1$ ) that imposes  $K = 1$ . This specification, however, still requires the simultaneous estimation of a large number of parameters if more than just a few variables are being jointly modeled. In particular, in a first-order model with  $p = 1$  and  $q = 1$ , there are  $n(n+1)$  free parameters of the second moment equations to be simultaneously estimated if variance targeting is employed. The number of parameters, therefore, still grows with the square of the number of variables. For a portfolio that includes 42 stocks, as in the empirical study presented in Chapter 6, this would require the estimation of 1806<sup>32</sup> parameters for the auto-regressive and lagged conditional terms of the second-moment equations. Diagonal GARCH models (see for example Bollerslev, Engle and Wooldridge (1988)) are a special case of BEKK( $p,q,1$ ) with  $A_1$  and  $B_1$  diagonal matrices that greatly economize on the number of parameters to be simultaneously estimated at the price of imposing a somewhat arbitrary restriction (i.e., all the cross-moments on the right-hand side of the same equation have the same coefficient).

When the number of variables and, therefore, parameters to be simultaneously estimated is high relative to the number of observations, the econometrician is typically left with few degrees of freedom in inferring the parameters values from the data. This makes inference procedure unreliable. The likelihood function is said to become ‘flat’ (its gradient is close to zero throughout the parameter space rather than just in the neighborhood of the optimum). Researchers have responded to this

---

<sup>32</sup> For example, with the number of assets  $n = 42$ , the number of parameters is  $2[n(n-1)/2+n] = 1,806$ .



difficulty in three ways. *First*, they have restricted the order of their multivariate GARCH specifications to no more than a few variables. *Second*, they have imposed restrictions on the correlation structure. Bollerslev (1990) proposes a class of multivariate GARCH (MVGARCH) models in which the conditional correlations are constant and thus the conditional covariances are proportional to the product of the corresponding conditional standard deviations. This restriction highly reduces the number of unknown parameters and thus simplifies estimation. It also guarantees positive-definiteness of the estimated variance-covariance matrices (if the unconditional correlation matrix is positive definite). This class of multivariate GARCH models is known as constant conditional correlation (CCC) GARCH. For example, in his multivariate GARCH analysis of European exchange rate volatility transmission, Bollerslev (1990) assumed constant conditional correlations to reduce the number of matrix inversions from 10,323 to 31 in estimating 30 parameters from 333 observations on 10 exchange rate returns. This approach has been commonly used in multivariate GARCH models of stock market volatility transmission (see, for example, Karolyi (1995), Koutmos and Booth (1995), Koutmos (1996), and Theodossiou, Kahya, Koutmos and Christofi (1997)). Although a computationally efficient solution, the assumption of constant correlations is becoming increasingly untenable, as further evidence emerges on the issue, see for example De Santis and Gerard (1997), Ang and Bekaert (2002) and Longin and Solnik (2001).

More recently, a *third* solution has been proposed to more efficiently model the time-series behaviour of large correlation matrices. This work includes Alexander's (2001, 2002) orthogonal GARCH (O-GARCH) model and the Dynamic Conditional Correlation GARCH (DCC-GARCH) model of Engle (2001) and Engle and Sheppard (2002). This family of models may be a useful way to describe the evolution over time of the correlation matrix of large systems. Orthogonal models are a special class of factor models. The latter models are based on the assumption that the data can be described as a linear transformation of a set of uncorrelated components. Orthogonal models require further that the linear transformation be represented by an orthogonal matrix. The latter is computed from the eigenvectors of



the data. The factors are chosen to be the principal components of the data or a subset of them (typically those with the highest explanatory power of the variation in the data, i.e. the highest coefficient of determination  $R^2$ ) and modelled as univariate zero mean ARCH type processes. The variances and the covariances of the assets are derived from the variances of the factors and the orthogonal linear transformation matrix. No idiosyncratic, non-factor related second moments need to be modelled because they are assumed not to influence the hypothesized multivariate conditional second moment process. If only the principal components of the system with the highest explanatory power are selected as factors, this leads to neglecting a large amount of non-systematic, temporary 'noise' in the data and, as a consequence, correlation estimates are usually more stable than with other more general methods.

While this approach has the advantage of simplicity and of generating stable correlation matrices, it has the drawback that the linear transformation used to orthogonalize the data and the parameters of the factor variance process are not estimated conditioning on the same information set (typically, the linear transformation matrix is computed orthogonalizing an unconditional variance-covariance matrix). This leads to an obvious loss of efficiency if the underlying factor structure is time-varying. Also, the generated variance-covariance matrices are not positive definite (they are positive semi definite). This might represent a problem in certain financial applications that require inversions of the variance-covariance matrix, such as portfolio optimization. Moreover this approach is best suited to model the variation in highly correlated systems such as domestic equity portfolios. These are systems where a few principal components explain most of the variation. Alexander (2001) reports encouraging estimation results for equity returns and exchange rates volatility and correlation estimates. However, Morillo and Pohlman (2002), using daily and weekly data on the 24 largest international stock market indices included in the MSCI World Index, find that O-GARCH variance-covariance matrix estimates do not improve on a simple EWMA univariate estimator.



Engle and Sheppard (2001) and Engle's (2002) DCC-GARCH model can be seen as a dynamic generalization of the CCC-GARCH. It allows for conditional time variations of the correlation matrix but imposes the restriction that the dynamics of all its elements are governed by common parameters. Moreover, the dynamics of the variance of each of the  $n$  variables being modelled do not interact with the dynamics of their correlations. These restrictions reduce dramatically the number of parameters to be estimated and ensure positive definiteness of the conditional correlation and variance-covariance matrices. The DCC-GARCH model has the additional advantage that it can be estimated using a two-stage procedure that helps further reduce the computational difficulties typical of multivariate GARCH models. In the first stage, univariate GARCH models are estimated for each asset, and the standard deviations estimates thus obtained are used to standardize the return innovations. In the second stage, a simple (usually scalar) specification is used to model the time-varying correlation matrix, which is obtained using the standardized residuals from the first stage. Engle and Sheppard (2001) show that this two-stage procedure yields consistent maximum likelihood parameter estimates. The inefficiency in the two-stage estimation process can be taken into account by modifying the asymptotic covariance of the correlation estimation parameters.

The DCC-GARCH model opens the door to more flexible variance-covariance matrices in the variance part. Since the conditional variances (together with the conditional means) can be estimated using univariate models, one can easily extend the DCC-GARCH by using more complex GARCH-type structures for the  $n$  univariate variance processes. Moreover, since the DCC-GARCH model, at least in the DCC(1,1) case, can be seen as a scalar BEKK(1,1,1) model for the standardized conditional innovations squares and cross-products, it also facilitates the specifications of richer dynamics of conditional correlations.

### 3.4.3 DCC-GARCH

In Chapter 6, I will use the DCC-MVGARCH model to estimate the conditional variance-covariance matrix of a large number of stocks and equity and bond indices. Therefore, it is now appropriate to provide a more exhaustive description of this model. I begin by rewriting, using (3.14) and (3.22), the multivariate log-likelihood of the observations on  $u_t$ :

$$\begin{aligned}
 L &= -0.5 \sum_{t=1}^T \left[ n \log(2\pi) + \log(|H_t|) + u_t' H_t^{-1} u_t \right] \\
 &= -0.5 \sum_{t=1}^T \left[ n \log(2\pi) + \log(|D_t C_t D_t|) + u_t' D_t^{-1} C_t^{-1} D_t^{-1} u_t \right] \\
 &= -0.5 \sum_{t=1}^T \left[ n \log(2\pi) + 2 \log(|D_t|) + \log(|C_t|) + \varepsilon_t' C_t^{-1} \varepsilon_t \right] \quad (3.39)
 \end{aligned}$$

Two components can vary in this likelihood function. The first part contains only terms in  $D_t$  and the second part contains only terms in  $C_t$ . Engle and Sheppard (2001) propose maximising  $L$  in two steps to overcome the well-known computational problems of MVGARCH models. They first maximise  $L$  with respect to the parameters that govern the process of  $D_t$ . This can be done by estimating univariate models of the returns on each stock nested within a univariate GARCH model of their conditional variance. One simple specification for the GARCH process followed by  $D_t^2$  is the following.

$$D_t^2 = \bar{D}^2 (1 - A - B) + A(u_{t-1} u_{t-1}') + B D_{t-1}^2 \quad (3.40)$$

Here,  $\bar{D}$ ,  $A$  and  $B$  are  $n \times n$  diagonal non-negative coefficient matrices that yield consistent, time-varying, estimates of  $D_t$ . The matrix  $\bar{D}$  is the long-run, baseline level to which the conditional volatilities mean-revert. To hasten the estimation procedure,  $\bar{D}$  can be set equal to the unconditional standard deviation matrix over



the sample (but in this case, the resulting conditional correlation matrix estimate is not guaranteed to be positive-definite). Engle and Sheppard (2001) suggest then maximising the second part of the likelihood function over the parameters of the process of  $C_t$ , conditional on the estimated  $D_t$ . This entails standardising  $u_t$  by the estimated  $D_t$  to obtain an  $nx1$  vector of standardized return innovations  $D_t^{-1}u_t$ . The maximum likelihood estimates of the parameters of the process of  $C_t$  that maximise the second part of (3.39) can then be found by estimating a multivariate model of  $D_t^{-1}u_t$ , nested within a multivariate scalar GARCH model of its conditional second moment matrix<sup>33</sup>. Ensuring positive-definiteness of this estimator is relatively easy to achieve as it simply requires, for the univariate part, using the same restrictions as univariate GARCH models and, for the multivariate part, using the standard quadratic form to allow for the estimation of the conditional correlation process without the need to impose lower bounds on its parameters. This is tantamount to treating the DCC model as a scalar BEKK model.

One simple specification for the GARCH process followed by  $C_t$  is the following:

$$C_t = \bar{C}(1 - \alpha - \beta) + \alpha \varepsilon_{t-1} \varepsilon'_{t-1} + \beta C_{t-1} \quad (3.41)$$

Here,  $\alpha$  and  $\beta$  are non-negative scalar matrices (all the elements on the main diagonal are equal)<sup>34</sup> and  $\bar{C}$  is a  $nxn$  matrix with all the elements along the main

<sup>33</sup> Notice that, in general, the conditional variance-covariance matrix of the first step estimates of  $D_t^{-1}u_t$ , is not  $C_t$  but its process is governed by the same parameters. In particular, defining  $\hat{D}_t$  as the first step volatility estimates from (3.40),  $C_t^* = E_{t-1}(\hat{D}_t^{-1}u_t u_t' \hat{D}_t^{-1})$  is the conditional variance-covariance matrix of the standardized innovations first step estimates. Engle (2002) proves that, when  $C_t$  is governed by the appropriate process and in particular when the latter can be described as a multivariate scalar GARCH model,  $C_t^*$  is governed by a process that has the same parameters as the process that governs  $C_t$ . Engle (2002) also proves that, given consistent estimates of  $C_t^*$ , consistent estimates of  $C_t$  can be retrieved as  $\hat{C}_t = \text{diag}(\hat{C}_t^*)^{-1} \hat{C}_t^* \text{diag}(\hat{C}_t^*)^{-1}$ , where  $\text{diag}(C_t^*)^{-1}$  is the matrix with the elements along the main diagonal equal to the main diagonal of  $C_t^{*-1}$  and the off-diagonal elements equal to zero. See Engle (2002) for details.

<sup>34</sup> Since  $\alpha$  and  $\beta$  are scalar matrices, to minimise the proliferation of symbols, I will denote the elements on their main diagonal with the same symbol as the matrices themselves.

diagonal equal to 1. The matrix  $\bar{C}$  is the long-run, baseline level to which the conditional correlations mean-revert. Again, to hasten the estimation procedure,  $\bar{C}$  can be set equal to the unconditional correlation matrix over the sample (but in this case the resulting conditional correlation matrix estimate is not guaranteed to be positive-definite). Obviously, more general specifications are possible. One such specification is the model proposed by Cappiello, Engle and Sheppard's (2003) that allow for asymmetric responses to past innovations of different sign and for a structural break date in the mean of the correlation process:

$$H_t \equiv D_t C_t D_t$$

Where:

$$\begin{aligned} D_t^2 &= \bar{D}^2 (1 - A - B) + A(u_{t-1}u'_{t-1}) + BD_{t-1}^2 \\ C_t &= \bar{C}(1 - \alpha - \beta) - \bar{S}\theta + \alpha\varepsilon_{t-1}\varepsilon'_{t-1} + \beta C_{t-1} + \theta S_{t-1} \end{aligned} \quad (3.42)$$

Here, the elements of the  $n \times n$  matrix  $S_{t-1}$  are the outer-products of 2 vectors that contain only negative return innovations,  $\bar{S}$  is the unconditional correlation matrix of the negative return innovations<sup>35</sup>,  $\theta$  is the coefficient of the matrix  $S_{t-1}$ . Notice that when the coefficient  $\theta$  is not constrained to be zero, the correlation process can be asymmetric. A symmetric DCC model gives higher tail dependence for both upper and lower tails of the multiperiod joint density. An asymmetric DCC gives higher tail dependence in the lower tail of the multi-period density. To complete the notation, the unconditional correlation matrix to which the correlation process is forced to mean-revert,  $\bar{C}$ , takes the value  $Q_1$  if  $t < \tau$  and  $Q_2$  if  $t > \tau$ , where  $\tau$  represents a selected structural break date and  $Q_1$  and  $Q_2$  are the sample average of  $\varepsilon_{t-1}\varepsilon'_{t-1}$  before and after the date  $\tau$ , respectively. Similarly,  $\bar{S}$  takes the value  $N_1$  if  $t < \tau$  and the value  $N_2$  if  $t > \tau$ , where  $N_1$  and  $N_2$  are the sample average of  $S_{t-1}$  before and after  $\tau$ , respectively.

---

<sup>35</sup> I estimate this using the sample average of the negative return innovation cross-products.



Since the 2-step estimation procedure yields consistent but inefficient maximum likelihood parameter estimates, a modification of the asymptotic covariance of the correlation estimation parameters is needed in order to obtain consistent standard errors of the parameter estimates. These depend on the cross-partial derivatives of the second stage likelihood with respect to the first and second stage parameters in addition to the typical Bollerslev and Wooldridge (1992) robust standard errors. This modification, however, is unlikely to be quantitatively very important.

In a recent application of the DCC-GARCH model to global equity and bond markets, Cappiello, Engle and Sheppard (2003) examine the correlation dynamics between equity markets in 21 countries and bond markets in 13 countries, using weekly data over the period from January 1987 to February 2001. These researchers show that the assumption of constant correlations can be dismissed in virtually all cases, and they proceed to model the time variation in the correlations while demonstrating the existence of strong persistence over time. Moreover, they find evidence that correlations react asymmetrically to joint return innovations of positive or negative sign.

#### *3.4.4 DCC-GARCH Extensions*

Using the DCC-GARCH specification renders the estimation of large variance-covariance matrices feasible since the number of parameters to be simultaneously estimated is considerably lower than in the case of more general multivariate specification and because the DCC-GARCH model can be estimated consistently in two stages. However, it is precisely when the number of assets to be jointly modelled is large that the restriction of common dynamics in the correlations is likely to be most unappealing. For example, as shown in Engle and Sheppard (2001), the DCC model leads to sub-optimal portfolio selection when the number of assets is large (typically above 20) because of the restriction that the asset-specific conditional correlations all follow the same dynamic structure. Hence, intuitively, when one considers many returns, one would want to allow for asset-specific

dynamics. Hafner and Franses (2003) generalize the DCC model by allowing one of the correlation parameters to vary across the assets. They label this specification generalized dynamic conditional correlation (GDCC) model. They estimate this model with daily data on 18 German stock returns, which are all included in the DAX, and for 25 UK stock returns in the FTSE. They find evidence that the GDCC model improves on both the CCC and the DCC specification. Billio, Caporin and Gobbo (2003) apply the restriction that correlations follow the same dynamics within groups of assets but allow for a richer behaviour of the correlations between groups. They call this specification block DCC (BDCC). Using data on Italian stocks from various industries, they estimate the variance-covariance matrix using a CCC, a DCC and a BDCC model. They use these estimates to form conditional minimum-variance portfolios. They show that the portfolios with, respectively, the highest and the lowest ex post variances are those selected using, respectively, the BDCC and the CCC conditional variance-covariance matrix estimates. The drawback of these generalizations, Hafner and Franses' (2003) GDCC and Billio, Caporin and Gobbo's (2003) BDCC, is that convergence becomes rather slow. Moreover, while it is relatively straightforward to extend the DCC model to allow for leverage effects in the correlation process, such extension would be more difficult in the GDCC and BDCC case.

### 3.4.5 *Financial Applications*

One of the main applications of second moment models and estimation techniques is the generation of forecasts for financial risk and portfolio management. Forecasts of asset volatilities and correlations, in particular, are required inputs for the estimation of portfolio Value at Risk (VaR), for portfolio optimization and for the construction of optimal hedge ratios. The VaR with significance level  $\alpha$  corresponds to the  $\alpha$ -th left quantile of the conditional distribution of portfolio returns (assuming that this has finite variance) over a given horizon (usually 10 days). The right quantile represents the VaR of a short position. In spite of its shortcomings, see Szego (2002) for a discussion, VaR is a widely used synthetic statistic for summarizing the level



of estimated risk exposure entailed by holding a portfolio. It was developed with the key contribution of the investment bank JP Morgan and it forms part of its RiskMetrics<sup>TM</sup> risk measurement methodology. This method is also used by financial regulatory authorities to set minimum capital requirement standards to protect the stability of the regulated financial intermediaries from market risk.

In a multivariate setting, the correlation matrix should be positive definite in order to ensure that correlations lie between -1 and 1 and that every sub-portfolio of assets under consideration has a correlation that lies between -1 and 1 with any other sub-portfolio. Imposing the further requirement that volatilities are non-negative ensures that the variance-covariance matrix is positive semi-definite. This is a desirable property of the estimated second moment variance-covariance matrices of asset returns as it ensures that the variance of every variable and of every combination of the variables is always non-negative. This rules out the possibility that investors can enjoy 'free-lunches' by forming risk-free positive-expected return arbitrage portfolios. To see this, assume instead that the variance-covariance matrix is not positive semi-definite. Therefore, there is some non-zero linear combination of assets (portfolios) with non positive variance. This represents an arbitrage opportunity whether it offers a positive, a negative or no expected return. If this combination of assets has positive expected return, it represents an arbitrage opportunity because the investor could expect to earn a positive return while being exposed to non positive variance ('negative variance' should not bother him). If the expected return is negative, it suffices to short sell the combination of assets to be in the same position as before, enjoying a positive expected return with non positive variance. If the expected return is zero, the investor can add (to the negative-variance combination of assets) portfolios with positive expected return and variance, up to the point where the overall variance is zero. At that point, the overall expected return will be by construction positive. Positive definiteness of the variance-covariance matrix is more restrictive and it ensures that the second moment matrix is invertible, thus making it possible to use it further in econometric (such as



a weighting matrix in weighted least squares regressions) and financial (notably in asset pricing and portfolio optimization algorithms) applications.

Because of its clustering behaviour, asset and especially stock price volatility typically exhibits high persistence (Engle and Patton (2001)), especially at relatively high frequencies (such as with weekly, daily and higher frequency data). It is therefore a natural candidate to be modelled using conditional autoregressive specifications such as ARCH and GARCH models. However, direct evidence on the extent to which ARCH type models display additional forecasting power relative to simpler unconditional alternatives is mixed. Akgiray (1989) finds that the forecast of US stock market volatility provided by a GARCH specification outperforms those formed using simpler ARCH, EWMA and unconditional estimates. Pagan and Schwert (1990) compare GARCH, EGARCH, a Markov switching regime model and three non parametric models for forecasting monthly US stock return volatility and they find that the conditional parametric specifications perform best (especially the EGARCH) but forecasting ability is nonetheless poor. Franses and van Diik (1996) find that non-linear and asymmetric specifications, such as the GJR asymmetric model, are unable to outperform simpler linear GARCH specifications in forecasting the weekly volatility of various European stock market indices. This result is interesting because it casts doubt on the importance of the asymmetric component in the second moment process of European stock market indices, at least at the weekly frequency<sup>36</sup>. There is evidence, however, that the asymmetric response to return innovations might be more important outside Europe. See, for example, the results reported by Brailsford and Staff (1996) on the better predictive ability of the GJR-GARCH for the Australian stock market volatility relative to models that do not allow for an asymmetric news impact curve<sup>37</sup>.

---

<sup>36</sup> Kearney and Poti (2003) report further (in sample) evidence on the weakness of the linear asymmetric response of European stock market index correlations to negative and positive innovations. I will discuss this in the empirical part of this thesis.

<sup>37</sup> See also the in sample results reported by Cappiello, Engle and Sheppard (2003) for 21 international stock market indices (including 9 European markets).



One of the advantages of univariate models is their ease of estimation relative to more complex multivariate specifications. Univariate EWMA and GARCH models are therefore routinely used in the financial industry to estimate the VaR of portfolios of assets over a given time horizon (see, for example, Bauwens, Laurent and Rombouts (2003)). The conditional distribution of the portfolio return, under the assumption of conditional multivariate normality, can be deduced from the estimated first and second moments. With a portfolio of many assets, however, the univariate model of the portfolio second moment must be re-estimated every time portfolio weights change. This can be a serious drawback if the portfolio contains large positions in financial instruments with non linear payoffs, with payoffs that depend on the correlation structure of asset returns and instruments that require time consuming numerical procedures for their pricing. This problem does not arise, or it is considerably milder, if estimates of the full variance-covariance matrix are available. The estimated multivariate distribution can be directly used to compute the portfolio variance and hence VaR for any set of asset weights. This is one of the advantages of a multivariate model.

Multivariate models however have the drawback of being computationally very intensive. As a consequence, a number of industry applications entailing large scale conditional variance-covariance matrices estimates rely on diagonal specifications. For example, in JP Morgan's RiskMetrics™ procedure, each element of the conditional variance-covariance matrix is estimated, using exponential smoothing, as a univariate EWMA and, as such, this approach corresponds to a diagonal IGARCH(1,1) model in which all of the intercepts in the conditional variance-covariance matrix are fixed at zero. Moreover, all the elements along the main diagonal of the  $A_1$  and  $B_1$  matrices take identical values,  $a$  and  $b$  respectively, with  $b = 1 - a$ . Thus, this model imposes on the diagonal IGARCH(1,1) model a scalar restriction. The use of the same smoothing parameter ( $b = 0.94$  in the latest release of the RiskMetrics™ technical document) facilitates the implementation and guarantees that the estimated conditional variance-covariance matrices are positive definite. Nonetheless, when viewed as a data generating process as opposed to a



“filter”, the RiskMetrics™ procedure is formally degenerate (Nelson (1990)). Alternatively, in the spirit of orthogonal factor models such as the O-GARCH, many financial institutions adopt the simplifying assumption that most of the variation of asset returns is generated by a limited number of common factors whereas the residual variation is attributable to purely idiosyncratic (and, hence, negligible) sources of variability. As CLMX (2001) and Kearney and Potì (2005b) show for the United States and the Euro area, respectively, this can be in many circumstances a somewhat heroic assumption, as the idiosyncratic portion is the main component of total volatility and the number of assets needed to diversify it away is large and tends to increase at times of market distress (I will present and discuss empirical evidence on these phenomena in Chapter 7). In the case of multi-country portfolios of equities, bonds or currencies, the assumption that returns are generated by a limited number of common factors is usually applied within each country (maybe with the United States or a World index as a common global factor). If estimates of international correlations are available, a multi-country variance-covariance matrix can then be constructed. This practice highlights the importance of modelling country correlations, and it also explains the wealth of empirical contributions in this area.

Engle and Sheppard (2001) use the DCC-GARCH model to estimate the conditional variance-covariance matrix of up to 100 assets represented by S&P Sector Indices and Dow Jones Industrial Average stocks and conduct specification tests using the JP Morgan’s RiskMetrics™ industry standard EWMA as a benchmark. They examine the performance of the model using three criteria: the standard deviation of portfolios where asset returns are standardized by the estimated portfolio conditional variance-covariance matrix (under the null of the DCC-GARCH these portfolios have unit variance), accuracy of VaR estimates with a pre-selected confidence level (evaluated against the actual frequency of losses in excess of the estimated VaR threshold) and relative forecasting performance with respect to JP Morgan’s RiskMetrics™ industry benchmark. They show that the DCC-GARCH estimator combines a strong performance in capturing important empirical features of the



conditional variances and covariances with ease of implementation. Morillo and Pohlman (2002) estimate the variance-covariance matrix of daily and weekly returns on the 24 largest international stock market indices included in the MSCI World Index using sample unconditional estimators and various conditional models. They use their variance-covariance matrix estimates in a portfolio optimization exercise, and report that the optimal portfolio based on DCC-GARCH estimates dominates the optimal portfolios based on all the other estimates. On the other hand, O-GARCH and CC-GARCH do not compare favourably with estimates produced by the simple univariate EWMA model.

Another application of models that provide estimates of the variance-covariance matrix of asset returns is the computation of the optimal hedge ratio. This is used to hedge portfolios of assets by taking an offsetting position in only one (or a few, in order to keep transaction costs low) liquid financial instruments, typically a futures contract. This is calculated as the ratio of future contracts to be sold to minimize the variance of the return on the overall portfolio, given pre-existing holdings. Under the appropriate assumptions (see Bauwens, Laurent and Rombouts (2003)), it is a function of the regression coefficient of the pre-existing portfolio against the hedge or, equivalently, of the conditional covariance between the hedge and the pre-existing portfolio and of the variance of the hedge. Brooks, Henry and Persaud (2002) suggest, using daily hedging, that a static hedge ratio computed using simple OLS regression estimates reduces the variance of the portfolio by 90 percent whereas using conditional time-varying BEKK-GARCH(1,1) estimates further improves this ratio by a mere 2 percent. Adding asymmetry effects provides no further improvement. Sephton (1993), however, finds that portfolio variance is minimized in sample using a time-varying optimal hedge ratio formed on the basis of conditional second moments estimates. Bera, Garcia and Roh (1997) find this to be the case also out of sample, especially for a VEC specification. However, the trading costs that it is necessary to incur in order to construct a time-varying optimal hedge may be substantially higher relative to a hedge strategy based on a static hedge ratio. Lien, Tse and Tsui (2002) find that, because of trading costs, a static

hedge strategy based on unconditional second moments estimates can be more profitable, for the same level of variance, than a strategy that uses a time-varying conditionally optimal hedge.

### **3.5. Summary and Conclusion**

Little theory is available to explain movements in, and especially, co-movements in volatility and higher moments. Their description is still the realm of reduced-form statistical models. In this Chapter, I have shown that these reduced form statistical representations can be estimated using either unconditional or conditional methodologies. I then discussed their relative merits and drawbacks and reviewed their main financial applications.

Unconditional estimation methodologies have the appealing feature of simplicity. This makes them useful tools for the study of the behaviour over time of second moments of asset returns when working with large sets of the latter, such as in applications to large portfolios of stocks. However, traditional unconditional procedures, such as that employed by CLMX (2001) based on computing volatilities and correlations first and then fitting regression models to study their time-series behaviour, do suffer from limitations. One of these problems is that, because of an *ad hoc* specification of the second moment process, this approach under-utilizes the available information. The resulting volatility and correlation estimates are therefore likely to be inefficient, and to exhibit an unsatisfactory performance in forecasting applications. The second problem arises because of the aggregation of data at a given frequency into lower frequency variance and correlation estimates. This leads to a potential small sample problem.

Parametric conditional estimation largely overcomes these problems while presenting its own limitations in terms of computational difficulties and model identification. These techniques appear most suited to relatively smaller scale applications, with a relatively small number of variables to be jointly modelled. A



partial exception is represented by the DCC-MVGARCH model. This specification restricts the dynamic behaviour of the correlation matrix, and it rules out interactions of the latter with variances. However, it substantially reduces the computational burden while it allows modelling the dynamics of both volatilities and correlations<sup>38</sup>.

Having now completed the literature review section of the thesis, in subsequent chapters I will study applications to novel data sets of some of the methodologies presented thus far. In doing so, I will extend and modify the extant methodologies whenever appropriate. In Chapter 4, I will use an autoregressive specification of the return process with conditional heteroskedasticity to estimate the amount of discount factor volatility required to explain the estimated conditional volatility of the S&P Composite Index over the period 1871-2003. In Chapter 5, I will derive the beta-pricing representation of the 3M-CAPM, I will estimate it using US industry index data and test whether the parameter estimates satisfy NS, RA and NIARA. In Chapter 6, after extending the DCC-MVGARCH model by allowing for a deterministic time trend, I will apply it to test for asymmetric reactions to news of different sign and for structural breaks in the conditional correlations of Euro-area equity indices and Government bonds. In Chapter 7, I will apply CLMX's (2001) variance decomposition methodology to a unique dataset of the volatilities and correlations of all the stocks traded in the countries of the Euro-area since 1974 and I will extend this methodology by introducing the notion of average correlation. I will discuss the properties of the latter and its relation to CLMX's (2001) variance components, and I will show how to construct an average correlation time series.

---

<sup>38</sup> However, given currently available computing power, it is not yet an easy task to apply the DCC-MVGARCH model to study the second moment dynamics of very large sets of assets. For example, using a Pentium 4 processor 2.40 GHz with 256 MB of Ram, I found that it takes roughly an hour, 3 and 5 hours to estimate a DCC-MVGARCH(1,1) specification with, respectively, 20, 30 and 42 assets. Above a threshold of approximately 50 assets, computing times increase exponentially and convergence becomes difficult to achieve with specifications that include asymmetry in the conditional correlation process.

## Chapter 4: The Cross-Section of Stock Returns

### 4.1. Introduction

As shown by Harvey and Siddique (2000) and Dittmar (2002), there is extensive evidence that US stocks exhibit systematic coskewness. In this Chapter, I will study the cross section of average excess returns on industry-sorted portfolios of US stocks and I will test whether, as predicted by the 3M-CAPM, there is a non-zero market price for systematic coskewness. A non-zero market price for systematic asset coskewness implies that investors are rewarded for holding assets that become more (less) volatile and correlated with their overall portfolio during market downturns (upturns). In checking this, I will focus on the unconditional implications of both conditional beta pricing models and conditional beta *and* gamma pricing models, or equivalently on stochastic discount factor models under quadratic and cubic approximation of investors' utility functions with possibly time varying shape.

In these models, while conditional betas and gammas, and in general the conditional moments and cross moments depend on the conditioning information set, conditional risk premia also change as a function of the parameters of the utility function, and hence of the conditioning information on which they depend. I will allow, however, only for time-variation in risk premia induced by variation in the shape of the utility function while I will not model variation in beta and gamma coefficients. This choice is motivated by the fact that realized betas (and gammas) display very little persistence even though their constituent realized volatilities and covariances are persistent. As a consequence, as argued by Andersen, Bollerslev, Diebold and Wu (2004), allowing for time-varying betas may do more harm than good and lead to spurious estimation results. This Chapter is based on two papers, Poti (2005b and 2005c), presented at the Annual Meeting of the Financial



Management Association (New Orleans, 2004) and discussed at the Doctoral Colloquium of the European Finance Association (Moscow, 2005).

In the next section, I derive the beta pricing representation of the 3-Moment conditional CAPM, henceforth 3M-(C)CAPM, formulated by Harvey and Siddique (2000). This derivation is, to my knowledge, novel. In Section 4.3, I present my dataset. In Section 4.4, I specify empirical versions of my beta pricing representation of the 2-moment (C)CAPM, henceforth 2M-(C)CAPM, and the 3M-CAPM. Within a 2-pass procedure, I estimate, in the first step, beta coefficients and, in the second step, the associated risk premia. I then check whether the estimated models satisfy the non satiation (NS), risk aversion (RA) and non increasing absolute risk aversion (NIARA) requirements. In Section 4.5, I discuss the consequences of a non concave utility function for the existence of a 3M-CAPM equilibrium and the implications of my 3M-CAPM parameter estimates. In Section 4.6, I contrast the 2M-(C)CAPM and the 3M-CAPM by estimating the correlation between their beta coefficients estimated over a rolling 30-month period. In Section 4.7, I estimate by GMM in a conditional setting the alternative beta-gamma representation of the 3M-CAPM proposed by Kraus and Litzenberger (1976). This way, I directly impose the NS, RA and NIARA requirements. In the final section, I summarize my findings and present my conclusions.

#### 4.2. The 3M-(C)CAPM Beta Pricing Representation

Using (2.20) to (2.22) and (2.60), I rewrite the conditional asset pricing equation in (2.59) as follows:

$$E_t(r_{i,t+1}) = \beta'_i \lambda_t \quad (4.1)$$

Where:

$$\beta_{i1,t} = \frac{Cov_t(r_{i,t+1}, r_{m,t+1})Var_t(r_{m,t+1}^2) - Skew_t(r_{m,t+1})Cov_t(r_{i,t+1}, r_{m,t+1}^2)}{Var_t(r_{m,t+1})Var_t(r_{m,t+1}^2) - [Skew_t(r_{m,t+1})]^2}$$

$$\beta_{i2,t} = \frac{Cov_t(r_{i,t+1}, r_{m,t+1}^2) Var_t(r_{m,t+1}) - Skew_t(r_{m,t+1}) Cov_t(r_{i,t+1}, r_{m,t+1})}{Var_t(r_{m,t+1}) Var_t(r_{m,t+1}^2) - [Skew_t(r_{m,t+1})]^2}$$

$$\lambda_t = - Var_t(f_{t+1}) b_t$$

$$= - \begin{bmatrix} Var_t(r_{m,t+1}) & Cov_t(r_{m,t+1}, r_{m,t+1}^2) \\ Cov_t(r_{m,t+1}, r_{m,t+1}^2) & Var_t(r_{m,t+1}^2) \end{bmatrix} \begin{bmatrix} b_{1,t} \\ b_{2,t} \end{bmatrix}$$

Equation (4.1) is a beta-pricing representation of Harvey and Siddique's (2000) conditional 3M-CAPM, or 3M-(C)CAPM. The factor loadings, the parameters  $\beta_{i1,t}$  and  $\beta_{i2,t}$ , are functions of the market variance, its skewness and the covariance and coskewness of the asset  $i$  with the market. They are both risk measures and coefficients of the regression of the asset excess-return on the market excess return and its square. From this point of view, they are analogous to the CAPM beta coefficient. The elements of the  $\lambda_t$  vector are the corresponding risk premia. Under RA,  $\lambda_{1t}$  should be positive to compensate investor for systematic covariance, whereas under NIARA  $\lambda_{2t}$  should be non-positive. In particular, under DIARA,  $\lambda_{2t}$  should be negative because investors should be willing to accept a lower average return to hold assets with positive coskewness.

Under the rational expectation assumption (RE), equation (4.1) is a testable restriction that the 3M-CAPM imposes on the cross section of expected (average) asset returns. The upshot of this formulation relative to the specification proposed by Harvey and Siddique (2000) is that I can test the unconditional implications of the conditional 3M-CAPM for the cross section of asset returns using a simple two pass estimation procedure, by regressing asset excess returns on the factors to obtain the factor loadings and the average excess returns on the estimated factor loadings to obtain the risk premia.



### 4.3. Data

I use the 30 Fama and French (1995) US industry portfolios, constructed using the stocks listed on the NYSE, NASDAQ and AMEX included in the Centre for Research on Security Prices (CRSP) database, as proxies for the investable universe of risky assets. As argued by Dittmar (2002), these industry portfolios represent a challenge for asset pricing models because they display considerable cross-sectional variation. As a proxy for the market portfolio of risky assets I use the market capitalization-weighted portfolio formed using the stocks included in the industry indices. I also use monthly and quarterly returns on 1-month and 3-month US Government Treasury Bills as proxy for the risk free rate. Finally, I use  $cay_t$ , the quarterly consumption-wealth ratio per capita estimates produced by Lettau and Ludvigson (2001), as a conditioning variable. This choice is motivated by the ability of  $cay_t$  to forecast future aggregate returns as reported by Lettau and Ludvigson (2001). The sample period is 1963-2002.

### 4.4. Beta-Pricing Empirical Specification

While Harvey and Siddique's (2000) 3M-CAPM implies the conditional stochastic discount factor model in (2.60), they estimate it without allowing for time variation in the parameters of the investor's marginal utility. Therefore, Harvey and Siddique (2000) effectively estimate an unconditional rather than a conditional 3M-CAPM, even though they do allow for time variation in betas. To formulate an empirical specification of the 3M-(C)CAPM beta-pricing representation in (4.1) that allows for time variation in the parameters of the investor's marginal utility, I let the  $a_t$  and  $b_t$  parameters of the stochastic discount factor vary as a linear function of the conditioning information provided by the variable  $z_t$ ,  $a = a^0 + a^1 z_t$ ,  $b_{1,t} = b_1^0 + b_1^1 z_t$ ,  $b_{2,t} = b_2^0 + b_2^1 z_t$ . Treating the vector  $f_{t+1} = [z_t \quad r_{m,t+1} \quad z_t r_{m,t+1} \quad r_{m,t+1}^2 \quad z_t r_{m,t+1}^2]'$  as the factors, this implies the following beta-pricing representation:

$$r_i = \alpha + \beta_{i3}\lambda_3 + \beta_{i4}\lambda_4 + \beta_{i5}\lambda_5 + \beta_{i6}\lambda_6 + \beta_{i7}\lambda_7 + \varepsilon_i \quad (4.2)$$

Here,  $r_i$  is the sample average of  $r_{it}$  and the elements of the  $\lambda$  vector are the cross-sectional parameter estimates of average excess returns on the corresponding elements of the  $\beta$  vector. The latter are the parameter estimates of the following time series regressions:

$$r_{it} = \alpha_i + \beta_{i3} z_{t-1} + \beta_{i4} r_{mt} + \beta_{i5} z_{t-1} r_{mt} + \beta_{i6} r_{mt}^2 + \beta_{i7} z_{t-1} r_{mt}^2 + \varepsilon_{it} \quad (4.3)$$

I allow for an intercept in (4.2) and (4.3). A model that is fully successful at explaining the cross section of excess returns should have  $\alpha = 0$ . Using  $cay_t$  as the conditioning variable, I set  $z_t = cay_t$  in (4.3). The model in (4.2) and (4.3) can be estimated by a 2-pass procedure that involves time series and cross-sectional regressions.

In the first pass, I estimate in a maximum likelihood setting the system of time-series regressions equations in (4.3) for the industries in my sample. I do not impose any constraint on the contemporaneous covariance of the residuals nor on their variance. I do correct, however, the variance and covariance matrix of the estimates for possible error autocorrelation and heteroskedasticity. The 3M-CAPM and 2M-(C)CAPM beta coefficients are reported in Table 4.1. They are the estimated  $\beta_{i4}$  and  $\beta_{i6}$ , and  $\beta_{i3}$ ,  $\beta_{i4}$  and  $\beta_{i5}$  respectively when all other beta coefficients are restricted to be zero. The 2M-CAPM and 3M-(C)CAPM beta coefficients are not reported to save space.

In the second pass of the estimation procedure, I then use my estimated beta coefficients as the regressors of average industry excess returns in a cross-sectional regression based on (4.2). The second pass estimation results are reported in Panel A of Table 4.2. The coefficient of determination  $R^2$  of the unrestricted 3M-(C)CAPM is surprisingly high, slightly above 31 percent (17 percent adjusted), for a model that does not include among the regressors portfolios returns that mimic additional partially ad-hoc factors such as size and the book to market ratio. The conditional



model however does not greatly improve on the unconditional 3M-CAPM. The  $R^2$  of the latter is slightly lower, almost 28 percent, but it is larger once we adjust for the degrees of freedom (22 percent). All the coefficient estimates of the unconditional 3M-CAPM are statistically significant. The (C)CAPM performs considerably worse than the 3M-(C)CAPM and the unconditional 3M-CAPM. Its  $R^2$  is just 7.5 percent (the adjusted one is marginally negative) and none of the coefficients estimates, with the exception of the intercept, are statistically different from zero at conventional significance levels. In Figure 4.1, I plot actual versus explained average industry portfolio returns. For some industries, such as Smoke, Books, Steel, Utilities, the model is off by more than 0.5 percent per quarter (roughly 2 percent per year). Other industries, however, such as Games, Constructions, Autos, Carry, Mines, Telecom, Paper and Wholesale, are priced in a remarkably accurate manner. These results provide evidence that systematic asset coskewness does help explain the cross-section of average returns. Even explicitly allowing for conditional time-variation in the shape of the utility function does not drive out its cross-sectional explanatory power.

Both the 3M-CAPM and the 3M-(C)CAPM estimates, however, imply a non concave shape of the utility function, that is incompatible with the risk aversion requirement. This can be seen by solving the risk premia equations in (4.1) for the parameters of  $m_{t+1}$ :

$$b_t = - \text{Var}_t(f_{t+1})^{-1} \lambda_t \quad (4.4)$$

I report in Figure 4.2 the stochastic discount factor  $m_{t+1}$  implied by the 2M-CCAPM, the 3M-CAPM and the 3M-(C)CAPM parameter estimates. These are locally consistent in all three cases with investors' non-satiation and preference for skewness. However, only the 2M-(C)CAPM stochastic discount factor displays risk aversion for every value taken by the market excess return over the sample. Both the 3M-CAPM and the 3M-(C)CAPM parameter estimates imply risk aversion only over excess returns below 1.5 percent. Above this threshold, the shape of the

estimated stochastic discount factor implies risk seeking. In other words, these estimates imply an inverse S-shaped utility function.



**Table 4.1**  
**3M-CAPM and (C)CAPM Regression Estimates**

<i>Portfolio</i>	$\beta_{i4}^{3M-CAPM}$	$\beta_{i6}^{3M-CAPM}$	$\beta_{i3}^{CCAPM}$	$\beta_{i4}^{CCAPM}$	$\beta_{i5}^{CCAPM}$
Food	0.9127 [23.00]	-0.2592 [-1.05]	0.7163 [ 2.33]	0.8843 [ 21.28]	-3.4764 [-1.06]
Beer	0.9015 [17.36]	0.095 [ 0.29]	-0.4512 [-1.12]	0.9186 [ 16.89]	6.7762 [ 1.58]
Smoke	0.739 [ 9.43]	0.3566 [ 0.73]	0.9272 [ 1.54]	0.6909 [ 8.47]	12.1417 [ 1.89]
Games	1.3525 [19.11]	-0.0566 [-0.13]	0.0435 [ 0.08]	1.3543 [ 18.25]	-1.9563 [-0.33]
Books	1.1958 [24.81]	-0.0398 [-0.13]	-0.4161 [-1.11]	1.2157 [ 23.93]	-2.2736 [-0.57]
Hshld	1.041 [25.63]	-0.2812 [-1.11]	0.2325 [ 0.73]	1.0335 [ 24.04]	-0.4501 [-0.13]
Clths	1.2632 [18.85]	0.0282 [ 0.07]	0.1613 [ 0.31]	1.2575 [ 17.71]	-2.6635 [-0.48]
Hlth	1.0209 [20.63]	-0.3191 [-1.03]	-0.3385 [-0.88]	1.0359 [ 19.88]	5.8596 [ 1.43]
Chems	0.9678 [24.78]	0.078 [ 0.32]	-0.004 [-0.01]	0.9636 [ 23.68]	8.0694 [ 2.52]
Txtls	1.1138 [16.78]	0.0788 [ 0.19]	-0.1176 [-0.23]	1.1143 [ 15.93]	7.8246 [ 1.42]
Cnstr	1.1981 [33.15]	0.1443 [ 0.64]	-0.4087 [-1.45]	1.2161 [ 31.94]	-1.4462 [-0.48]
Steel	0.9495 [16.71]	0.3355 [ 0.95]	-0.7223 [-1.65]	0.9746 [ 16.43]	10.7442 [ 2.30]
FabPr	1.1475 [27.81]	-0.0744 [-0.29]	-0.2149 [-0.67]	1.1587 [ 26.57]	-2.0503 [-0.60]
ElcEq	1.1347 [23.45]	-0.0199 [-0.07]	0.0087 [ 0.02]	1.1355 [ 22.20]	-3.4563 [-0.86]
Autos	1.0398 [17.04]	0.351 [ 0.92]	-0.296 [-0.62]	1.0501 [ 16.22]	0.0861 [ 0.02]
Carry	1.1798 [18.81]	-0.3035 [-0.78]	0.6857 [ 1.40]	1.151 [ 17.39]	-0.3157 [-0.06]
Mines	0.8484 [10.65]	0.6626 [ 1.33]	-0.4745 [-0.76]	0.8648 [ 10.20]	1.7157 [ 0.26]
Coal	0.9636 [11.09]	0.4109 [ 0.76]	0.1942 [ 0.29]	0.9488 [ 10.34]	9.3328 [ 1.29]
Oil	0.7283 [13.18]	0.1914 [ 0.56]	0.5048 [ 1.17]	0.7061 [ 12.09]	-2.0122 [-0.44]
Util	0.6114 [14.35]	-0.0072 [-0.03]	0.5194 [ 1.57]	0.5866 [ 13.16]	4.0633 [ 1.16]
Telcm	0.7135 [15.82]	0.2758 [ 0.98]	-0.2389 [-0.68]	0.7203 [ 15.09]	4.2731 [ 1.14]
Servs	1.2758 [22.94]	0.2336 [ 0.67]	-0.4588 [-1.06]	1.2967 [ 22.10]	-3.3338 [-0.72]
BusEq	1.1804 [23.24]	0.192 [ 0.61]	-0.7322 [-1.85]	1.2098 [ 22.66]	2.881 [ 0.69]
Paper	0.9854 [27.18]	-0.0373 [-0.16]	-0.3134 [-1.11]	0.9989 [ 26.08]	1.4261 [ 0.47]
Trans	1.1667 [24.56]	-0.2128 [-0.72]	0.0819 [ 0.22]	1.1625 [ 23.11]	2.7965 [ 0.71]
Whlsl	1.2018 [26.37]	-0.0448 [-0.16]	-0.6947 [-1.98]	1.2348 [ 25.99]	-3.9501 [-1.06]
Rtail	1.142 [23.85]	0.1406 [ 0.47]	-0.618 [-1.66]	1.1674 [ 23.15]	1.6503 [ 0.42]
Meals	1.4072 [20.96]	-0.0707 [-0.17]	-0.5342 [-1.02]	1.4322 [ 20.19]	-1.5632 [-0.28]
Fin	1.0869 [34.56]	-0.229 [-1.17]	0.2066 [ 0.84]	1.0799 [ 32.37]	-2.0016 [-0.76]
Other	1.1318 [27.67]	-0.4124 [-1.62]	0.279 [ 0.87]	1.1226 [ 25.85]	0.2948 [ 0.09]

*Notes.* This table reports the parameters estimates and significance t-statistics for the time series regression models based on the 3M-CAPM and on the (C)CAPM. All the symbols are defined as in the text.

**Table 4.2**  
**3M-CAPM, CCAPM and 3M-CCAPM**

*Panel A*  
(3S-LS Regression Estimates)

	$\alpha$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\lambda_6$	$\lambda_7$	$R^2$ (Adj. $R^2$ )
3M- CCAPM	1.06% [.019]	-.07% [.714]	.80% [.069]	-.01% [.091]	-.54% [.000]	.003% [.009]	31.17% (16.83%)
3M- CAPM	1.06% [.013]		.89% [.023]		-.50% [.000]		27.87% (22.52%)
3M- CCAPM	1.22% [.010]	.30% [.233]	.72% [.105]	.00% [.581]			7.49% (-3.18%)

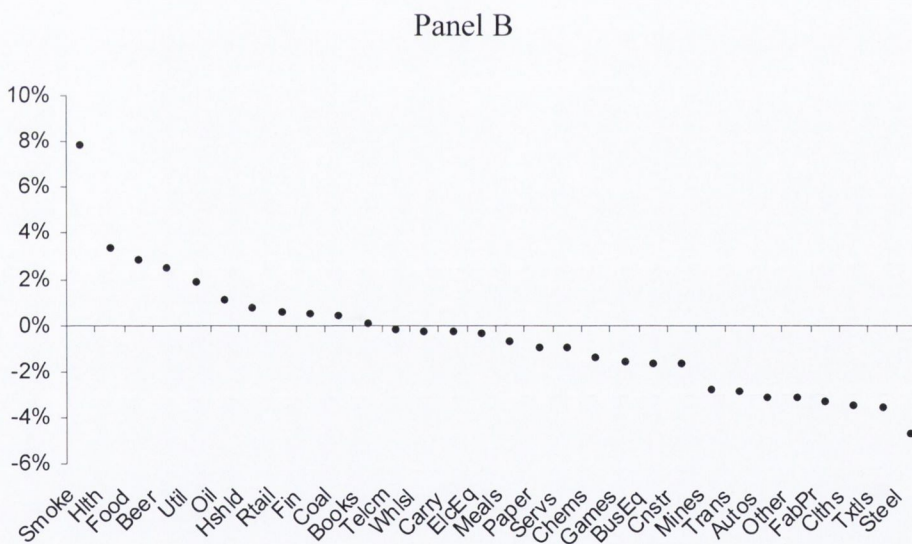
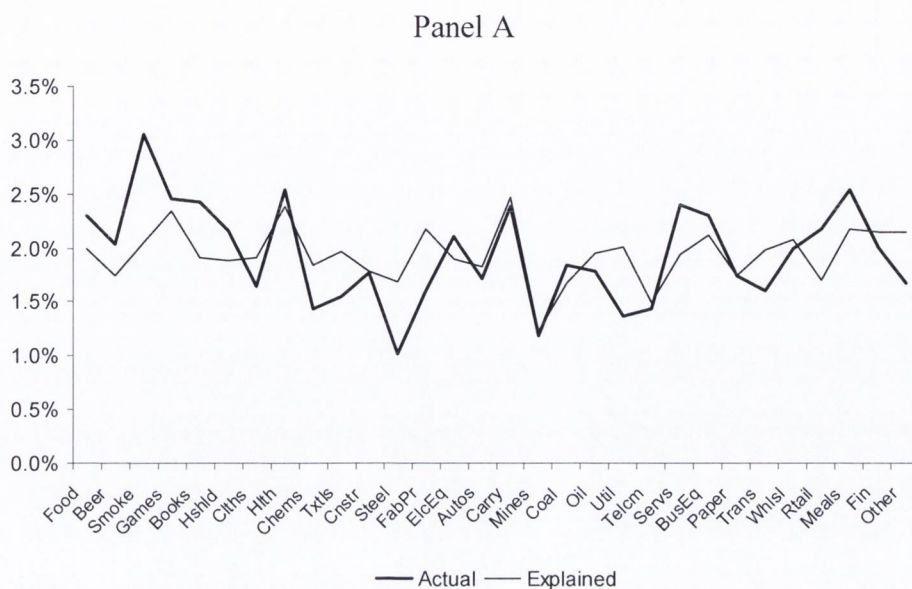
*Panel B*  
(GMM Estimates)

Model	Constr.	Estimat.	DF	TJ	$\theta_4$	$\theta_5$	$\theta_6$	$\theta_7$	$\delta_1$ (%)	$\delta_2$ (%)
3M- CAPM		GMM	28	32.48 [.255]	-1.26 [.000]		-7.66 [.000]		6.95	-1.3
3M- CAPM	RA + NLARA	GMM	28	51.90 [.003]	-1.08 [.000]		.000		5.26	0.0
3M- CCAPM		GMM	26	34.34 [.126]	-1.08 [.000]	-33.01 [.321]	-8.28 [.000]	114.8 [.322]	9.99	-4.0
3M- CCAPM	RA + NLARA	GMM	26	60.55 [.000]	-1.14 [.000]	-106.9 [.000]	0.00 [.000]	277.0 [.013]	5.43	0.0

*Notes.* Panel A of this table reports measures of fit ( $R^2$  and adjusted  $R^2$ ) for the model that nests both the 3M-CAPM and the CCAPM, for the 3M-CAPM and for the CCAPM. Panel B reports the GMM estimation results for various set of orthogonality conditions that correspond to the 3M-CAPM (top two rows) and to the 3M-CCAPM that nests both the 3M-CAPM and the CCAPM (lower two rows). *DF* denotes degrees of freedom (number of orthogonality conditions in excess of the number of parameters to be estimated). The expression *TJ* is *T* (the sample size) times Hansen's (1982) *J* statistic and it is distributed as a Chi-Squared with degrees of freedom equal to the number of over-identifying restrictions (*DF*). All the other variables are defined as in the text. The risk premia  $\delta$  are annualised. Significance levels of t-statistics appear in brackets. The sample period is 1963-2002.

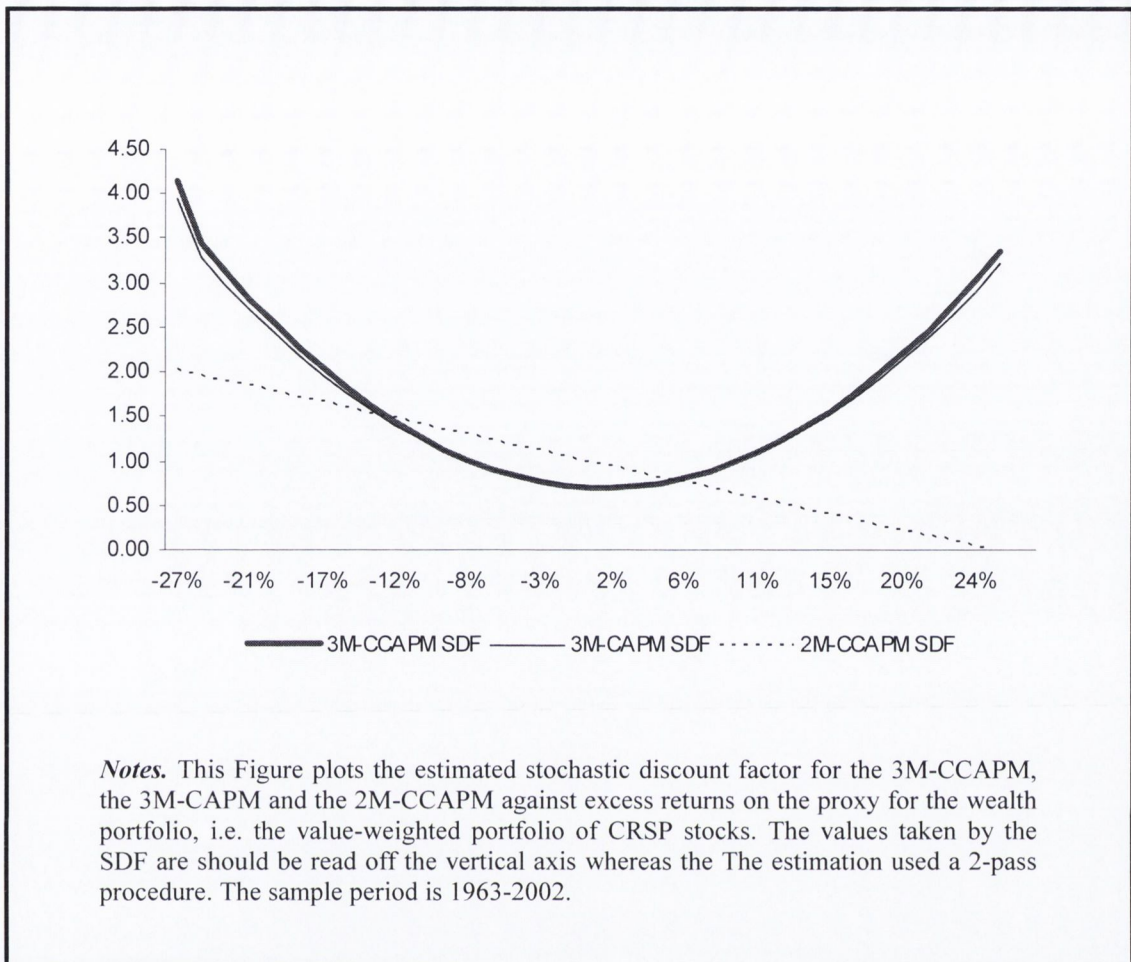


**Figure 4.1**  
**Actual vs. Explained Average Excess Returns**  
**3M-(C)APM**



*Notes.* The darker line in Panel A plots the quarterly actual and explained (by the 3M-(C)CAPM) average excess return for the industry portfolios reported along the horizontal axis. The estimation used a 2-step procedure with system OLS estimates for the first step. Panel B plots the annualized unexpected returns sorted in ascending order. The sample period is 1963-2002.

**Figure 4.2**  
**Estimated SDF**



#### 4.5. Concavity of the Utility Function and the 3M-CAPM

As acknowledged by Post, Levy and Van Vliet (2003), concavity of the utility function is not a law of nature. There is evidence that investors display local risk seeking behaviour (Kahneman and Tversky (1979), Statman (2002)) and there are numerous non-standard theories that allow for alternative assumptions about the shape of the utility function, see Shefrin (2005) for a review. In particular, Shefrin and Statman (2002) Behavioural Portfolio Theory predicts risk aversion over losses



and risk seeking over gains and thus an inverse S-shaped utility function. This shape is also advocated by Post and Levy (2002).

If we accept some risk seeking over gains and an inverse S-shaped utility function, with risk aversion for losses and risk seeking for gains, my beta pricing representation of the 3M-(C)CAPM is surprisingly successful at explaining the cross section of industry returns (once we take into account that it does not include any ad hoc factor), with a coefficient of determination in the region of 30 percent. Thus, Markowitz (1952) type utility functions, that capture investors' twin desire for downside protection in bear markets and upside potential in bull markets, can explain a large portion of the cross-section of stock returns.

If the utility function is non-concave, however, it is not guaranteed that the market portfolio is efficient. This is because, in a constrained optimization problem, a stationary point is guaranteed to represent a maximum only when the objective function is quasi-concave but only a concave utility function can guarantee that expected utility is quasi-concave. The reason for this is the mathematical fact that the sum of concave functions is guaranteed to be quasi-concave<sup>40</sup> whereas the sum of quasi-concave functions is not guaranteed to be quasi-concave. Lacking the quasi-concavity requirement, the parameters of the stochastic discount factor that satisfy the first order conditions  $0 = E_t[m_{t+1}(\theta)r_{i,t+1}]$  are not guaranteed to represent the constrained maximum of the expected utility function. Thus, there is no guarantee that orthogonality conditions such as (2.1) and (2.50) represent first order conditions for the maximization of a utility function. Similarly, there is no guarantee that the stochastic discount factor parameters are such that the market portfolio maximizes expected utility, even though the first order conditions  $0 = E_t(m_{t+1}r_{i,t+1})$  are satisfied. A similar conclusion was reached by Post, Levy and Van Vliet (2003), but they erroneously argue that a stationary point is guaranteed to represent a maximum only for concave (rather than quasi-concave) expected utility functions.

---

<sup>40</sup> This is not the case for concave functions. The sum of concave functions is guaranteed to be concave. This is why the concavity of utility guarantees the concavity of the expected utility function.

#### 4.6. The 3M-CAPM vs. the (C)CAPM

As discussed in Section 2.5.10, some of the empirical features of asset excess returns predicted by the 3M-CAPM coskewness premium can in principle be explained by a conditional asset pricing model with time-varying risk aversion, if the square of the market return and hence its second moments proxy for the state of the wider economy (e.g. distress and recessions).

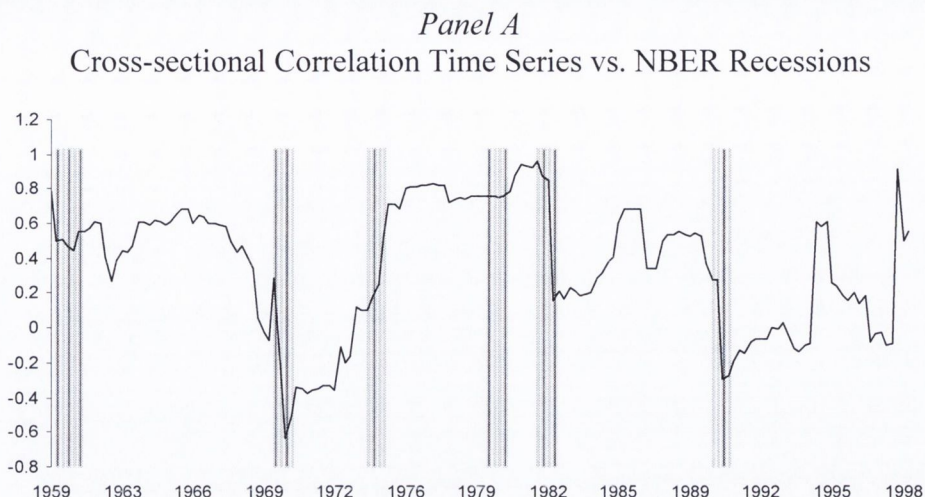
How could one 'check' on this? As a first check, I compute the cross-sectional correlation between the 2M-CCAPM  $\beta_{i5}$  coefficients, estimated setting  $\beta_{i6}$  and  $\beta_{i7}$  equal to zero, and the 3M-CAPM  $\beta_{i6}$  coefficients, estimated setting  $\beta_{i3}$ ,  $\beta_{i5}$  and  $\beta_{i7}$  equal to zero. Their linear cross-sectional correlation coefficient is 0.38. The positive value of the correlation coefficients agrees with the intuition that the  $\beta_{i5}$  and the  $\beta_{i6}$  capture a similar type of exposure and that asset co-skewness, therefore, 'picks up' the cyclical changes in the CAPM beta that occur as we move from expansion to recession.

Next, I redo the analysis with a 30-month rolling window. Visual inspection of Figure 4.3 provides further hints on the relation between the  $\beta_{i5}$  and the  $\beta_{i6}$  coefficients. Panel A of this Figure plots their cross-sectional 30-month rolling correlation against NBER (National Bureau for Economic Research) recessions while Panel B plots their cross-sectional 30-month rolling correlation against 30-month rolling market returns. While, from Panel A, this correlation is high at the beginning and low at the end of 3 out of 6 reported recessions, it is Panel B that provides the most suggestive clues. With the exception of the period 1959-1972, the plots of the cross-sectional correlation and of the market returns series seem to share the same 'trend' (with some sort of lag structure and a lot of noise) in the 3 (roughly) decades 1972-1982, 1982-1992, 1993-1999. This suggests that, while volatility exposure (captured by coskewness and thus  $\beta_{i6}$ ) is almost the same as conditional beta exposure (i.e. the sensitivity of asset betas to the state of the economy captured



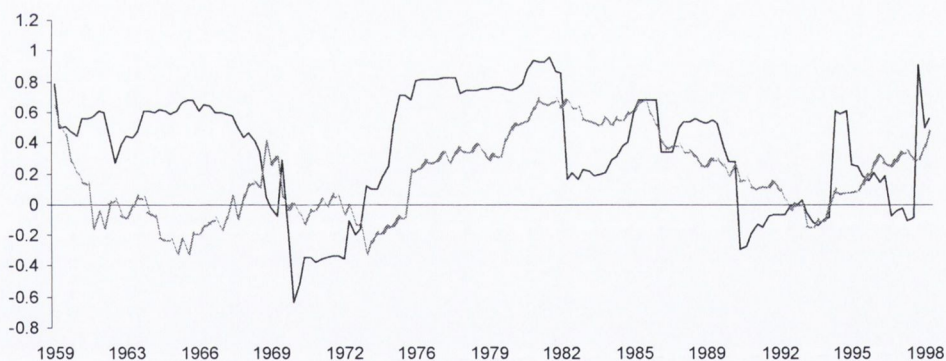
by  $\beta_{is}$ ) at the peak of bull markets, at the bottom of a bear market they are almost unrelated. Since adding coskewness increases the explanatory power of the cross-section of average industry returns more than adding the conditioning variable, the former likely captures features about investors' attitude towards risky assets in the middle of a market downturn that the latter misses.

**Figure 4.3**  
**Gamma Coefficients vs. Return Sensitivities**



*Panel B*  
 Cross-sectional Correlation Time Series vs. Market Portfolio Returns

— Cross-correlation — Market Returns



*Notes.* The darker line in this figure plots the 30-quarter (7 years and 6 months) rolling cross-sectional correlation between the 2M-(C)CAPM coefficients  $\beta_{15}$  (multiplied by a factor 10 to facilitate visual comparison) and the 3M-CAPM coefficients  $\beta_{16}$ . The former are the sensitivities of industry excess returns to the product of the market excess-return and the conditioning variable (the lagged consumption-wealth ratio) whereas the latter are the coskewness coefficients. In Panel A the shaded areas represent NBER recessions. In Panel B the jagged line represents the rolling time series of 30-quarter returns on the market portfolio.



#### 4.7. GMM Estimates with RA and NIARA

To impose the RA and NIARA conditions, I generalize the beta-gamma representation of the 3M-CAPM proposed by Kraus and Litzenberger (1976) by allowing for conditional time variation in the parameters of the SDF, and I estimate it by GMM. In the conditional version of the Kraus and Litzenberger (1976) 3-MCAPM, henceforth 3M-(C)CAPM, I account for time-variation in the parameters of the utility function by letting them depend in a linear fashion on a conditioning variable  $z$  that represents the available information set:

$$\begin{aligned}\theta_1 &= \theta_4 + \theta_5 z \\ \theta_2 &= \theta_6 + \theta_7 z\end{aligned}\tag{4.5}$$

For convenience, I reproduce the Kraus and Litzenberger's (1976) beta-gamma representation of the 3M-CAPM already introduced in (2.54):

$$E(r_i) \cong \delta_1 \beta_i + \delta_2 \gamma_i$$

Where,

$$\begin{aligned}\delta_1 &= \frac{-E[u''(r_m | \theta)]E[r_m - E(r_m)]^2}{E[u'(r_m | \theta)]} \\ \delta_2 &= \frac{-\frac{1}{2}E[u'''(r_m | \theta)]E[r_m - E(r_m)]^3}{E[u'(r_m | \theta)]} \\ \beta_i &= \frac{E[(r_i - E(r_i))(r_m - E(r_m))]}{E[r_m - E(r_m)]^2} \\ \gamma_i &= \frac{E[(r_i - E(r_i))(r_m - E(r_m))^2]}{E[r_m - E(r_m)]^3}\end{aligned}$$

Plugging (4.5) into (2.54), I then recover the conditional risk premia  $\delta_1$  and  $\delta_2$  from the conditional utility function parameters as follows:

$$\delta_1 = \frac{-E[u''(r_m | \theta)]E[r_m - E(r_m)]^2}{E[u'(r_m | \theta)]} - \frac{[2\theta_4 + 2\theta_5 E(z) + 6\theta_6 E(r_m) + 6\theta_7 E(zr_m)]E[(r_m - E(r_m))]^2}{1 + 2\theta_4 E(r_m) + 2\theta_5 E(zr_m) + 3\theta_6 E(r_m^2) + 3\theta_7 E(zr_m^2)} \quad (4.6)$$

$$\delta_2 = \frac{-0.5E[u'''(r_m | \theta)]E[r_m - E(r_m)]^3}{E[u'(r_m | \theta)]} - \frac{[3\theta_6 + 3\theta_7 E(z)]E[(r_m - E(r_m))]^3}{1 + 2\theta_4 E(r_m) + 2\theta_5 E(zr_m) + 3\theta_6 E(r_m^2) + 3\theta_7 E(zr_m^2)} \quad (4.7)$$

The above equations give the unconditional asset pricing implications in terms of the two risk factors, the market excess return and its square, of a conditional specification of the 3M-CAPM Euler equations in (2.50). Under non satiation, marginal utility must always be positive. Therefore the denominator of (4.7) is always positive. Since variance is always positive, risk aversion (RA) and hence  $u''(r_m | \theta) \leq 0$  imply  $\delta_1 \geq 0$ . Similarly, since the skewness of the market portfolio is empirically found to be negative, preference for skewness and hence NIARA and  $u'''(r_m | \theta) \geq 0$  imply a positive  $\delta_2$ , i.e.  $\delta_2 \geq 0$ . RA in turn implies:

$$u''(r_m) = 2\theta_4 + 2\theta_5 z + 6(\theta_6 + \theta_7 z) r_m \leq 0 \quad (4.8)$$

A cubic utility function cannot be concave over its entire domain, thus the condition in (4.8) should hold only over the sample values of  $r_m$  and  $z$ . This condition is in general difficult to impose. However, when NIARA holds:

$$u'''(r_m) = 6(\theta_6 + \theta_7 z) \geq 0 \quad (4.9)$$

In this case, a sufficient condition for (4.8) and thus for RA to hold over the sample values of  $r_m$  and  $z$  is the following:

$$2\theta_4 + 2\theta_5 z + 6[\theta_6 + \theta_7 \text{Max}(z)] \text{Max}(r_m) \leq 0 \quad (4.10)$$



Here, the operator  $Max()$  denotes the sample maximum of the argument and (4.9) and (4.10) together are sufficient conditions for RA to hold over the relevant portion of the representative investor's utility function domain. Finally, the following condition adapted from Post, Levy and van Vliet (2003) constraints the market premium to be the sum of the beta and gamma premium:

$$\begin{aligned} E[u'(r_m)r_i]w &= E[(1 + 2\theta_4 r_m + 2\theta_5 z r_m + 3\theta_6 r_m^2 + 3\theta_7 z r_m^2)r_i]w \\ &= E(r_m) + 2\theta_4 E(r_m^2) + 2\theta_5 E(z^2 r_m^2) + 3\theta_6 E(r_m^3) + 3\theta_7 E(z^3 r_m^3) = 0 \end{aligned} \quad (4.11)$$

Here,  $w$  represents the vector of value-weights (which can be in principle time-varying). In the empirical applications, I replace the unconditional expectations in (4.5) to (4.11) by the corresponding sample moments and I estimate them by GMM. Imposing (2.26) on the excess-returns on the 30 Fama and French (1995) US industry portfolios and the CRSP market portfolio and using (4.5) yields a system of 30 orthogonality conditions. I estimate both the unconditional and the conditional model. The unconditional model imposes on the conditional one the restriction that, in (4.5),  $\theta_5 = 0$  and  $\theta_7 = 0$ . I implement the GMM methodology by estimating this system of orthogonality conditions with the constraint in (4.11), with and without the constraints in (4.9) and (4.10), by multivariate non-linear system least squares. I then construct first stage GMM standard errors with the identity matrix used as the weighting matrix for the orthogonality conditions. I use the identity matrix instead of Hansen's (1982) optimal weighting matrix because I place more importance on robustness than on efficiency. As discussed by Cochrane (2001), the optimal weighting matrix places more importance on moment conditions that are more precisely estimated and thus on the industry indices with less volatile pricing errors. The identity matrix instead places the same weight on all moment conditions and thus on all industries. Since in my study the pricing errors on all the industries are equally important, the identity is a preferable choice. An important difference between the two pass regression and GMM is that the latter, in the unconstrained

estimation, forces the intercept to take a zero value. The 2-pass regression is therefore a more robust estimation procedure.

The empirical results are reported in Panel B of Table 4.2. Here, the coefficient  $\delta_1$  is the beta premium and the coefficient  $\delta_2$  is the gamma premium. As explained in Section 2.6.1, these are the coefficients a beta-gamma representation of the implications of (2.50) for the cross-section of asset returns. These estimates are different from the estimates of the beta-pricing representation presented in Panel A of Table 4.2. It is worth recalling that the assumption of greed implies  $E[u'(r_m | \theta)] > 0$  and RA implies  $E[u''(r_m | \theta)] \leq 0$ . Thus, since  $E[r_m - E(r_m)]^2 \geq 0$ , the beta premium is *positive* for risk-averse, greedy investors, i.e. in (2.54)

$$\delta_1 = \frac{-E[u''(r_m | \theta)]E[r_m - E(r_m)]^2}{E[u'(r_m | \theta)]} > 0. \text{ If the market portfolio skewness is negative}$$

(as it is often empirically the case),  $E[r_m - E(r_m)]^3 \leq 0$ . Also, under the assumption of NIARA,  $E[u'''(r_m | \theta)] \geq 0$ . As a consequence, the gamma premium

$$\delta_2 = \frac{-\frac{1}{2}E[u'''(r_m | \theta)]E[r_m - E(r_m)]^3}{E[u'(r_m | \theta)]} \text{ is } \textit{positive} \text{ for greedy investors with decreasing}$$

absolute risk aversion. Thus, while the beta premia under the beta-pricing and beta-gamma representations,  $\lambda_4$  and  $\delta_1$  respectively, have the same sign, the coskewness  $\lambda_6$  and gamma premia  $\delta_2$  have opposite sign.

When I impose the RA and NIARA constraints, the gamma premium becomes close to zero and both the conditional and the unconditional model are rejected at the 0.3 percent significance level by Hansen's (1982)  $J$  test. This confirms the result reported by Dittmar (2002) and by Post, Levy and van Vliet (2003). It is interesting to note, however, that it is the NIARA restriction that is binding ( $\theta_6 = 0$ ), whereas the RA constraint is slack in a neighbourhood of the stationary point (the proxy for the market portfolio). In other words, the only way to allow for a non-zero gamma premium is to allow for a negative third derivative of utility, i.e.  $u'''(r_m) < 0$ . This, however, implies from (4.7) a negative gamma premium, skewness aversion instead



of skewness preference and an S-shaped utility function. It would also be interesting to impose NS and NIARA without imposing RA to replicate the results of the 2-pass regression. This is virtually impossible, however, as it is very difficult to impose NS without first imposing RA and imposing the latter without first imposing NIARA.

#### **4.8. Summary and Conclusions**

In this Chapter, I have derived the beta-pricing representation of Harvey and Siddique's (2000) specification of the 3M-CAPM. Relative to Harvey and Siddique (2000), the main innovation of this study is an explicitly conditional empirical specification of the stochastic discount factor and the derivation of the beta-pricing representation of their model. My beta pricing representation of Harvey and Siddique (2000) 3M-CCAPM is also different from the Kraus and Litzenberger (1976) beta-gamma representation since their beta and gamma are not regression coefficients.

From an empirical point of view, I have updated the evidence provided by Harvey and Siddique (2000) and by Dittmar (2002) on the ability of the coskewness and gamma premia to explain the cross-section of US industry returns. My sample spans almost 40 years from 1963 to 2002, whereas the sample period of the studies of Harvey and Siddique (2000) and Dittmar (2002) stops, respectively, in 1993 and 1995. I have also employed Lettau and Ludvigson's (2001) consumption-wealth ratio,  $cay_t$ , as a conditioning variable to model time variation in the parameters of the utility function. While there is considerable evidence that this variable can predict long horizon returns and thus proxy for time variation in the shape of the utility function, no study to date had previously used it as a conditioning variable in a 3 moment asset pricing model. In particular, it had not been used by Dittmar (2002).

I found that, while the 3M-CAPM is rejected when risk aversion and non increasing absolute risk aversion are imposed, its factor loadings are surprisingly successful at explaining the cross section of industry returns, with a coefficient of determination

between 20 and 30 percent. These values are high for a model that does not include among the regressors portfolios returns that mimic additional and partially ad-hoc factors such as size and the book to market ratio. Thus, while the beta pricing representation of a quadratic market factor model like (4.1) is unlikely the empirical specification of the 3M-CAPM, it does captures a substantial portion of the cross sectional variation of returns. It can therefore be interpreted as a statistical representation of asset return, but its theoretical motivation remains to be established. In other words, a quadratic specification that allows for both systematic asset covariance and coskewness to explain the cross section of returns should be seen as a relatively successful reduced form representation of a yet unknown theoretical asset pricing model.

In this reduced form representation, first moments (mean stock excess returns) are related to systematic second moment realizations (squared market returns). In particular, stocks that display high coskewness with the stock market and thus a high correlation with market volatility display lower average returns. Conversely, stocks with negative coskewness and thus with low correlation with market volatility display higher average returns. More specifically, for a given level of covariance with the stock market, stocks with low coskewness offer higher average returns than stocks with high coskewness.



## Chapter 5: European Equity Returns Second Moments

### 5.1. Introduction

In this Chapter, I extend the study of Campbell, Lettau, Malkiel and Xu (2001), henceforth CLMX (2001), to the European equity market. In particular, I analyse the behaviour over the period 1974-2004 of systematic and aggregate firm and industry-level volatility and correlation of the 3515 stocks listed on the markets of the current members of the EMU<sup>41</sup>. I focus on relatively low frequency time variation in second moments because it is more relevant from a strategic asset allocation perspective than high frequency (e.g. daily) movements. Thus I work with semi-annual volatilities and correlations estimates constructed from weekly returns. I study both their long-term trends and the shorter run relationships that link these series to each other and to aggregate returns. My aim is to extend the literature by applying CLMX (2001) methodology to a new dataset.

This Chapter is based on two studies, Kearney and Poti (2003) and Kearney and Poti (2005b). These have been presented at the European Finance Association (EFA) Annual Meeting (Glasgow, 2003), at the European Financial Management Association (EFMA) Annual Meeting (Milan, 2005) and Kearney and Poti (2005b) has been accepted for presentation at the Financial Management Association (FMA) Annual Meeting (Chicago, 2005).

I begin by introducing, in the next Section, a decomposition of average stock variance into systematic and idiosyncratic components. This is similar to CLMX's (2001) decomposition but it is based on returns instead of excess returns and I derive

---

<sup>41</sup> In this study I neglect the country level, traditionally prominent in the literature on volatility in European markets, see for example Baele (2002), and I focus instead on the firm, industry and aggregate level of the EMU stock market as a whole. This choice is motivated by the considerable evidence on a substantial degree of equity market integration, which has gathered pace in Europe since the mid-1990s (Hardouvelis, Malliaropoulos and Priestley (2000) and Fratzschler (2002)). Moreover, following the introduction of the Euro, equity markets of the countries that have adopted the new currency have become almost perfectly correlated, as reported by Cappiello, Engle and Sheppard (2003) and by Kearney and Poti (2003, 2005a).

it following an alternative, more intuitive approach. In Section 5.3, I then describe my data set and construct my variance and correlation series. In Section 5.4, I examine their long-run behaviour. In Section 5.5, I compare them to analogous series constructed from United States data. In Section 5.6, I discuss possible explanations for the observed long-run trends in individual stocks volatilities and correlations. Then, in Section 5.7, I study the lead-lag relations between the variance series and their ability to predict aggregate returns. In the final Section, I summarise the main findings and present some concluding remarks.

## 5.2. Variance Decomposition

Denote by  $R_{i,t}$  the return on asset  $i$  included in portfolio  $P$ . It can be decomposed into the conditionally risk free rate,  $R_{f,t}$ , a portfolio-related component and an asset-specific component:

$$R_{i,t} = R_{f,t} + \beta_{i,p}(R_{p,t} - R_{f,t}) + u_{i,t} \quad (5.1)$$

Here,  $R_{p,t}$  is the return on the portfolio  $P$ ,  $\beta_{i,p}$  is a regression coefficient and  $u_{i,t}$  is an idiosyncratic regression residual<sup>42</sup>. The unconditional variance of the asset can also be decomposed into a systematic and an idiosyncratic component:

$$\text{Var}(R_{i,t}) = (1 - \beta_{i,p})^2 \text{Var}(R_{f,t}) + \beta_{i,p}^2 \text{Var}(R_{p,t}) + \text{Var}(u_{i,t}) \quad (5.2)$$

Averaging across the assets, the variance of the typical asset can also be approximately decomposed into a systematic and an idiosyncratic component:

---

<sup>42</sup> Notice that idiosyncratic residuals are not assumed to be uncorrelated across all pair of firms and industries (our reference model is the CAPM, not the APT). They are, however, orthogonal on average. In other words, since they are regressions residuals of models that include the same set of regressors, their average correlation is by construction zero.



$$\begin{aligned} \text{Avg}[Var(R_{i,t})] &= \text{Avg}[(1 - \beta_{i,p})^2 Var(R_{f,t})] + \text{Avg}[\beta_{i,p}^2 Var(R_{p,t})] + \text{Avg}[Var(u_{i,t})] \\ &= \text{Avg}[(1 - \beta_{i,p})^2] Var(R_{f,t}) + \text{Avg}(\beta_{i,p}^2) Var(R_{p,t}) + \text{Avg}[Var(u_{i,t})] \end{aligned} \quad (5.3)$$

Here, the operator  $\text{Avg}(\cdot)$  denotes a weighted average across all the assets included in the portfolio. Using an elementary statistical result, and assuming that the cross-sectional variation of the beta coefficients,  $CSV(\beta_{i,p})$ , is not too high<sup>43</sup>,  $\text{Avg}(\beta_{i,p}^2)$  and  $\text{Avg}[(1 - \beta_{i,p})^2]$  in (5.3) can be conveniently approximated as follows:

$$\begin{aligned} \text{Avg}(\beta_{i,p}^2) &= \text{Avg}(\beta_{i,p}) \text{Avg}(\beta_{i,p}) + CSV(\beta_{i,p}) = 1 + CSV(\beta_{i,p}) \cong 1 \\ \text{Avg}[(1 - \beta_{i,p})^2] &= [\text{Avg}(1 - \beta_{i,p})]^2 + CSV(1 - \beta_{i,p}) = CSV(1 - \beta_{i,p}) \cong 0 \end{aligned} \quad (5.4)$$

Using (5.4), the decomposition of the variance of the typical asset in (5.3) collapses into the sum of the portfolio variance and of the average idiosyncratic variance:

$$\text{Avg}[Var(R_{i,t})] \cong Var(R_{p,t}) + \text{Avg}[Var(u_{i,t})] \quad (5.5)$$

Turning to a larger scale analysis, the returns on the industry indices and on the individual stocks in the market portfolio are described in equations (5.6) and (5.7).

$$R_{j,t} = R_{f,t} + \beta_{j,m} (R_{m,t} - R_{f,t}) + \varepsilon_{j,t} \quad (5.6)$$

$$R_{ij,t} = R_{f,t} + \beta_{ij,m} (R_{m,t} - R_{f,t}) + \beta_{ij,j} \varepsilon_{j,t} + e_{ij,t} \quad (5.7)$$

Here,  $R_{j,t}$  is the industry  $j$  return,  $R_{ij,t}$  is the return on firm  $i$  in industry  $j$ ,  $R_{m,t}$  is the return on the market portfolio,  $\beta_{j,m}$ ,  $\beta_{ij,m}$  and  $\beta_{ij,j}$ , are regression coefficients and  $\varepsilon_{j,t}$

<sup>43</sup> At least in the value weighted case, large stocks will likely have betas close to 1, i.e. betas close to their cross-sectional average. This means that the cross-sectional variation of betas in the market capitalization weighted (value-weighted) case must be low. Using United States data, CLMX (2001) show that the cross-sectional variance of beta coefficients is relatively small for large portfolios of stocks. They also demonstrate that the cross-sectional variance of beta coefficients is typically not very volatile over time and thus, in time series analysis, it can be safely ignored.

and  $e_{ij,t}$  are, respectively, industry and firm-level idiosyncratic regression residuals<sup>44</sup>.

Letting  $u_{ij,t} = \beta_{ij,j} \varepsilon_{j,t} + e_{ij,t}$ , (5.7) can be rewritten as follows:

$$R_{ij,t} = R_{f,t} + \beta_{ij,m} (R_{m,t} - R_{f,t}) + u_{ij,t} \quad (5.8)$$

By construction,  $R_{m,t}$ ,  $e_{ij,t}$ , and  $\varepsilon_{j,t}$  are orthogonal,  $u_{ij,t}$  is orthogonal to  $R_{m,t}$  and thus it is an idiosyncratic regression residual, so (5.8) decomposes returns into a pure market component and a pure idiosyncratic component and (5.7) decomposes the latter into pure industry and firm level components<sup>45</sup>.

Based on this model of returns and on (5.5), total stock variance can be decomposed into a systematic and an idiosyncratic component,

$$VAR_t \cong MKT_t + IDIO_t \quad (5.9)$$

where,

$$VAR_t = \sum_{j=1}^n w_{j,t} \sum_{i=1}^k w_{ij,t} Var(R_{ij,t})$$

$$MKT_t = Var(R_{m,t}) = \sum_{j=1}^n \sum_{i=1}^k w_{ij,t}^2 Var(R_{ij,t}) + \sum_{j=1}^n \sum_{i=1}^k w_{ij,t} Cov(R_{ij,t}, R_{m,t})$$

$$IDIO_t = \sum_{j=1}^n w_{j,t} \sum_{i=1}^k w_{ij,t} Var(u_{ij,t})$$

Here,  $k$  denotes the maximum number of stocks in each of the  $n$  industries,  $w_{j,t}$  the weight of industry  $j$  in portfolio  $m$ , and  $w_{ij,t}$  the weight of stock  $i$  in industry  $j$ ,  $VAR_t$  is the weighted average total stock variance,  $MKT_t$  is the variance of the market portfolio and  $IDIO_t$  is the average idiosyncratic variance. Intuitively,  $VAR_t$  can be interpreted as the variance of the typical stock, and  $IDIO_t$  as the variance borne by

<sup>44</sup> Notice that idiosyncratic residuals are not assumed to be uncorrelated across all pair of firms and industries (our reference model is the CAPM, not the APT). They are, however, orthogonal on average. In other words, since they are regressions residuals of models that include the same set of regressors, their average correlation is by construction zero.

<sup>45</sup> Moreover, letting  $\beta_{ij,m} = \beta_{ij,j} \beta_{j,m}$  and substituting from (7.6) into (7.7),  $R_{ij,t} = R_{f,t} + \beta_{ij,j} (R_{j,t} - R_{f,t}) + e_{ij,t}$ .



the arbitrageur that holds a long position in the typical stock and a short position in the market portfolio.

Since this framework can be applied to any portfolio, we can apply it to decompose the variance of the typical industry into its market and idiosyncratic components as follows:

$$VAR_t^{ind} \cong MKT_t + IND_t \quad (5.10)$$

where,

$$VAR_t^{ind} = \sum_{j=1}^n w_{j,t} Var(R_{j,t})$$

$$IND_t = \sum_{j=1}^n w_{j,t} Var(\varepsilon_{j,t})$$

Here,  $VAR_t^{ind}$  is average total industry variance and  $IND_t$  is the industry level average idiosyncratic variance. Intuitively, the former can be seen as the variance of the typical industry and the latter as the variance born by the typical arbitrageur that holds a market neutral long-short position in an industry index. The idiosyncratic portion of average total variance can then be further decomposed into its industry and firm level components:

$$IDIO_t = IND_t + FIRM_t \quad (5.11)$$

where,

$$\begin{aligned} FIRM_t &= IDIO_t - IND_t \\ &= \sum_{j=1}^n w_{j,t} \left[ \sum_{i=1}^k w_{ij,t} Var(u_{ij,t}) - Var(\varepsilon_{j,t}) \right] \\ &= \sum_{j=1}^n w_{j,t} \sum_{i=1}^k w_{ij,t} Var(e_{ij,t}) \end{aligned}$$

The last approximate equality follows from the application of (5.5) to  $u_{ij,t}$  in (5.8).  $FIRM_t$  is the firm-level average idiosyncratic variance. Intuitively, it can be

interpreted as the variance borne by the typical arbitrageur that holds a long position in the typical stock and a short position in the industry to which it belongs.

Since  $u_{ij,t}$  can be seen as a CAPM idiosyncratic residual, (5.9) and (5.11) provide a CAPM-equivalent decomposition<sup>46</sup> of average total variance into market variance and average idiosyncratic variance and its industry and firm components with the considerable advantage that it bypasses the need to estimate possibly time-varying *betas*. This variance decomposition is very similar to the decomposition proposed by CLMX (2001), but it is based on returns instead of excess returns and therefore it does not require the identification of the risk free rate. This represents a considerable advantage when it is unclear which rate constitutes the appropriate proxy for the risk-free return. Such a situation typically occurs when working with European returns in the pre-Euro period. Moreover, my derivation of the variance decomposition is complementary to the strategy followed by CLMX (2001) in that it sheds light on different intuitions about the relation between the systematic and idiosyncratic components of the volatility of the typical asset. It is also easier to generalize to any portfolio of assets and orthogonal portions thereof.

Furthermore, define the average volatility of the stocks included in the market portfolio as  $VOL_t = \sum_{j=1}^n w_{j,t} \sum_{i=1}^k w_{ij,t} \sqrt{Var(R_{ij,t})}$ . Assuming that the market portfolio is well diversified, the average stock correlation can be obtained as the ratio of the market variance to the square of the average stock volatility:

$$CORR_t = \frac{MKT_t}{VOL_t^2} \quad (5.12)$$

---

<sup>46</sup> As discussed by CLMX (2001), this is an approximate decomposition. In particular,  $IDIO_t$  is only approximately equal to the average variance of the CAPM idiosyncratic residuals. CLMX (2001), however, show that their difference is negligible if the cross-sectional variance of the beta coefficients is not too volatile.



Here,  $CORR_t$  is the level of correlation that, if assumed to hold for all pairs of stocks, would give the same market volatility as the full correlation matrix. In a similar way, defining  $VOL_t^{ind} = \sum_{j=1}^n w_{j,t} \sqrt{Var(R_{j,t})}$  as the average industry volatility, we can also construct a measure of average correlation for a diversified portfolio of industries as follows:

$$CORR_t^{ind} = \frac{MKT_t}{VOL_t^{ind}{}^2} \quad (5.13)$$

Equations (5.12) and (5.13) are based on a general and intuitively appealing result that, as proven in Appendix D, applies to any well diversified portfolio and that can therefore be used to simplify the construction of the average correlation time series of a large number of assets. This can be particularly useful for risk managers and derivatives traders. It also has the interesting analytical implication that the variance of a diversified portfolio can be modelled in either a univariate or a simple bivariate setting by studying the process followed by its average correlation, its average volatility, and their interaction. In Appendix E, to illustrate the usefulness of (5.12) and (5.13), I outline a dynamic trading strategy based on average correlation.

### 5.3. EMU Data and Variable Construction

I use market, industry and firm-level weekly returns and semi-annual capitalization data from *Datastream International Ltd.* for the period December 1974 to March 2004. By using weekly returns I overcome the problem of asynchronous trading across the EMU stock markets. Firm level data comprises the total returns and market capitalisation for the 3,515 stocks listed over the period 1974-2004 on the equity markets of Austria, Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Norway and Spain. These are the countries that had adopted the Euro as of March 2004. The industry level data is obtained from *Datastream International Ltd.* Level 4 fixed history industry indices for the Euro

area equity market<sup>47</sup>. The market data comprises total returns on the *Datastream International Ltd* fixed history<sup>48</sup> index for the overall Euro area stock market<sup>49</sup>.

Similarly to CLMX (2001), I employ unconditional estimators of variances based on sums (or averages) of return innovations squares and cross-products. This choice is motivated by the large number of stocks in my sample that renders multivariate conditional estimation methodologies unfeasible. Also, since I focus on low frequency estimates, conditional serial dependence in second moments is unlikely to be important. The implicit assumption of the unconditional estimation methodology is that the variance of a process is observable, and as pointed out by Merton (1980), it can be estimated to any desired degree of accuracy by sampling the squared deviations of the process realisations from their means at sufficiently high frequency. I therefore define variance over a period  $T$  of length  $p$  as the average of the squared deviations of returns (or their components)  $R_t$ ,  $t = 1, \dots, p$ , from their mean  $\bar{R}_T$ . In all computations I apply the convention that each year comprises 52 weeks and each semester comprises 26 weeks. Therefore, to compute my semi-annual variance of weekly returns, I set  $p$  equal to 26. Formally:

$$Var(R_t)_T = \sum_{t=1}^{26} (R_t - \bar{R}_T)^2 \quad (5.14)$$

Using (5.14), I first construct variance series using non-overlapping semi-annual periods for the individual stocks  $Var(R_{ij,t})_T$ , for the individual industries  $Var(R_{j,t})_T$

---

<sup>47</sup> Datastream Level 4 Industry Indices classify Euro area stocks into 35 industries (Panel A in Table 7.1), thus providing enough cross-sectional variation to be able to discriminate industry-specific variation from sources of variation common to all the stocks (e.g. the market).

<sup>48</sup> The choice of using fixed history indices is necessary to ensure consistency with my average variance computation methodology and with the procedure followed by CLMX (2001).

<sup>49</sup> I constructed a value-weighted index of all the stocks included in our dataset for the shorter period 1<sup>st</sup> semester 1997 – 1<sup>st</sup> semester 2004 and found that its correlation with the Datastream Euro area market index is almost perfect (96.8 percent) over this period and over various sub-periods. Thus, given the availability of the excellent proxy represented by the Datastream Euro area market index (that represents at least 75% of the capitalization of the Euro area equity market), I felt that it was not necessary to construct the value-weighted index of all the stocks for the entire 1974-2004 sample period, a computationally very intensive task that would have likely lead to errors.



and for the market portfolio  $Var(R_{m,t})_T$ . I then set  $MKT_T$  equal to the latter, compute the average total variance time series,  $VAR_T$ , and using (5.9) I construct the average idiosyncratic variance series,  $IDIO_T$ , as the difference between  $VAR_T$  and  $MKT_T$ . Turning to the decomposition of average idiosyncratic variance into its industry and firm level components, I use (5.10) to construct  $IND_T$  by subtracting  $MKT_T$  from  $VAR^{ind}_T$  and, using (5.11), I derive  $FIRM_T$  by subtracting  $IND_T$  from  $IDIO_T$ . Finally, applying (5.12) and (5.13) and using the constructed market, stock and industry variance series, I compute the average correlation among the stocks and industries.

This gives 61 non overlapping semi-annual variance and correlation data points ( $T = 1, 2, \dots, 61$ ) computed from the weekly returns data. The variance series are annualized by multiplying by 2 to minimise rounding errors and to display the results in a more intuitive form. While I construct both equally-weighted and value-weighted series, I focus mostly on the latter<sup>50</sup>. The constructed series and the notation employed are summarized in Panel B of Table 5.1. The decomposition of the value-weighted average total stock variance series into its market and idiosyncratic components is reported in Panel A of Figure 5.1, Panel B plots the ratio of firm to industry variance and Panel C reports average stock and industry correlations.

Inspection of Figure 5.1 reveals that total, idiosyncratic and market variance start off relatively low and tend to rise towards the end of the period. However, this tendency is more pronounced for idiosyncratic variance and its firm-level component. Idiosyncratic variance is the largest component of average total variance and average stock correlation is usually well below 50 percent with the noticeable exception of the 1974 oil crisis and the 1987 stock market crash. The potential benefit to diversification strategies is therefore substantial. These developments are broadly in line with those reported by CLMX (2001) for United States stocks.

---

<sup>50</sup> Equally-weighted series are available upon request.

**Table 5.1**  
**Data and Variables Definitions**

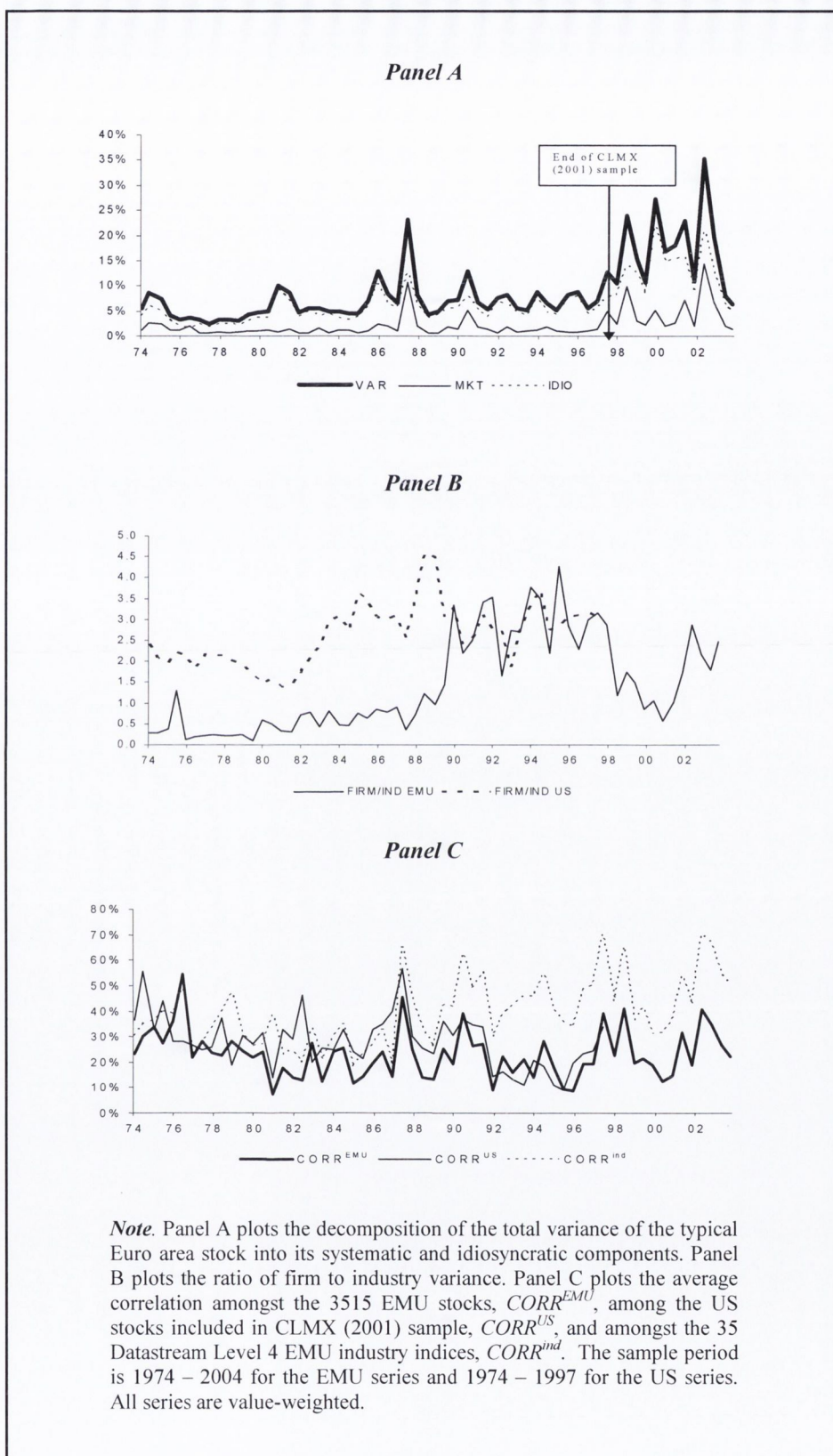
Panel A			
<i>Industries – Datastream Level 4</i>			
1	Mining	19	Retail, General
2	Oil & Gas	20	Leisure & Hotels
3	Chemicals	21	Media, Entertainment
4	Cons.. & Bldg. Mat.	22	Support Services
5	Forestry & Paper	23	Transport
6	Steel &Oth. Metals	24	Food & Drug Retailers
7	Aerospace, Defence	25	Telecom Services
8	Diversified Industrials	26	Electricity
9	Electric Equipment	27	Other Utilities
10	Eng.&Machinery	28	Inf. Tech. Hardw.
11	Auto & Parts	29	Software & Comp. Serv.
12	H'Hld GDS&Textls	30	Banks
13	Beverages	31	Insurance
14	Food PrDr./PrCr.	32	Life Assurance
15	Health	33	Investment Cos.
16	Per.Care&Hshld	34	Real Estate
17	Pharm. & Biotech	35	Spc. & Other Finance
18	Tobacco		

Panel B		
<i>Variables</i>		
1	$R_{jT}$	Weekly return on industry j
2	$R_{i,jT}$	Weekly return on stock i from industry j
2	$R_{mT}$	Weekly return on the stock market portfolio
3	$VAR_T$	Average total variance of stock returns
4	$MKT_T$	Annualised semi-annual variance of $R_{mT}$
5	$IDIO_T$	$VAR_T - MKT_T$
6	$VAR_T^{ind}$	Average total variance of industry returns
7	$IND_T$	$VAR_T^{ind} - MKT_T$
6	$FIRM_T$	$VAR_T - VAR_T^{ind}$

*Notes.* Panel A of this table reports the industries included in the sample based on the Datastream Level 4 classification. Panel B summarizes the main variables. The market portfolio is the Datastream index for the Euro area. All returns are total returns (they include accrued dividends). All indices are “fixed history” (they are not recalculated following modifications to the index composition).



**Figure 5.1**  
**Variance and Correlations Series**



#### 5.4. Time Series Behavior of Unconditional Second Moments

I begin my formal time series analysis by providing descriptive correlations and autocorrelations and by testing for the presence of long run trends. I then examine the short run interactions between my decomposed variance series. Table 5.2 presents descriptive correlations of the market variance, the average idiosyncratic variance and the average correlation series with each other's lags and with lags of industry and firm volatility, average correlation, the market excess-return and Gross Domestic Product (GDP) growth. The low persistence of the market variance and correlation series is due to their construction from relatively low frequency (weekly) returns and to the semi-annual sampling period and it suggests that they are unlikely to contain a unit root. This is also the case for the more persistent average idiosyncratic variance and its industry and firm-level components<sup>51</sup>. I therefore treat the constructed variance and correlation series as stationary and work in levels without differencing. All series display a negative correlation with stock market returns. They are also positively correlated with GDP growth and hence pro-cyclical.

##### *Long Run Trends*

To test for the presence of a deterministic time trend I estimate a dynamic model that includes among the regressors a constant and a lag of the dependent variable<sup>52</sup>. I then conduct a Wald-type test of the restriction that the deterministic time trend coefficient is zero. The results are reported in Table 5.3. The trend coefficient is significant in both the average idiosyncratic variance,  $IDIO_T$ , and the market

---

<sup>51</sup> While they are more auto-correlated, they appear far from containing a unit root. To double check on whether the series are stationary, however, I also conduct Dickey-Fuller and augmented Dickey-Fuller tests and I analyse the spectral density function of the series. These results are available upon request.

<sup>52</sup> I include among the regressors only one lag of the regressand because, from Table 7.2, higher order auto-correlations do not appear to be important. To check that the estimated residuals from this model are serially independent I use the Durbin's  $h$  statistic because, in the presence of a lagged value of the dependent variable among the regressors, the  $DW$  test is biased towards acceptance of the null of no autocorrelation. I use the generalised version of Durbin's  $h$ -test, developed by Godfrey and Breusch, based on a general Lagrange Multiplier test. Even though this procedure can detect higher order serial correlation, I test only the null of no first-order residual autocorrelation.



variance,  $MKT_T$  series. These trends explain a substantial portion of the rise in these series over time. After 5 years, for example, the projected increase in market variance,  $MKT_T$ , and in idiosyncratic variance,  $IDIO_T$ , is 0.56 percent and 1.0 percent respectively. These values correspond to increases in market volatility and idiosyncratic volatility of the typical stock of about 5.5 and 10 percent respectively. Since the time trend is statistically insignificant for average industry-level idiosyncratic variance,  $IND_T$ , but highly significant for firm-level idiosyncratic variance,  $FIRM_T$ , the surge in idiosyncratic variance,  $IDIO_T$ , is attributable mostly to an upward trend in firm-level volatilities. The upward trend in idiosyncratic and firm-level variance is similar in magnitude to the upward trend in the corresponding United States series studied by CLMX (2001). Market volatility, however, is not trended upwards in the CLMX (2001) sample.

The long run mean of average stock correlations,  $CORR_T$ , is close to 20 percent. The typical coefficient of determination  $R^2$  and hence the explanatory power of the market model, with zero intercept, is therefore rather low at about 4 percent (calculated as the square of 20 percent). The trend coefficient of average stock correlations,  $CORR_T$ , is not statistically significant. This is not surprising, given that both market variance,  $MKT_T$ , and idiosyncratic variance,  $IDIO_T$ , are trended upwards by similar magnitudes. A consequence of this finding is that, in the long run, the explanatory power of the market portfolio is relatively stable. CLMX (2001) results imply instead a downward trend in average stock correlation and in the explanatory power of the market portfolio in the United States.

**Table 5.2**  
**Descriptive Correlations**

	$q$	$MKT_{T+q}$	$IDIO_{T+q}$	$IND_{T+q}$	$FIRM_{T+q}$	$CORR_{T+q}$	$(R_m - R_f)_{T+q}$	$GDP_{T+q}$
$MKT_T$	1	0.22	0.32	0.14	0.38	0.10	0.04	-0.05
	0	1	0.74	0.55	0.70	0.61	-0.52	0.14
	-1	0.22	0.35	0.19	0.39	-0.06	-0.08	-0.08
	-2	0.24	0.35	0.23	0.35	0.08	-0.11	0.27
	-3	0.18	0.50	0.50	0.36	-0.12	0.10	0.10
$IDIO_T$	1	0.35	0.62	0.45	0.59	-0.01	-0.19	0.01
	0	0.74	1	0.81	0.88	0.08	-0.35	0.16
	-1	0.32	0.62	0.45	0.59	-0.14	-0.10	0.12
	-2	0.23	0.54	0.37	0.53	-0.11	0.03	0.29
	-3	0.34	0.60	0.49	0.53	-0.03	-0.02	0.18
$CORR_T$	1	-0.06	-0.14	-0.19	-0.07	0.18	0.19	0.03
	0	0.61	0.08	0.07	0.06	1	-0.30	0.20
	-1	0.10	-0.01	-0.06	0.04	0.18	-0.01	-0.15
	-2	0.08	-0.03	-0.08	0.01	0.23	-0.03	0.14
	-3	0.00	0.12	0.20	0.03	0.00	0.33	0.01

*Note.* This table reports descriptive correlations of the variables reported in the first column with leads  $q$  of the variables reported at the top of the other columns. No series is linearly de-trended. The proxy for the risk free rate is the semi-annual average of the 1 Month Euro-Mark.  $GDP$  is the GDP growth rate. The sample period is 1974-2004



**Table 5.3**  
**Long Run Trends**

$y_T$	$\alpha$ ( <i>t-stat.</i> )	$\delta$ ( <i>t-stat.</i> )	$\beta$ ( <i>t-stat.</i> )	<i>h-stat.</i> ( <i>sign.</i> )	<i>Wald-stat.</i> ( <i>sign.</i> )
<i>IDIO</i> <sub><i>T</i></sub>	1.16 (1.29)	0.10 (3.09)	35.86 (2.80)	0.52 (.470)	9.55 (.003)
<i>FIRM</i> <sub><i>T</i></sub>	-0.40 (0.88)	0.11 (4.98)	12.79 (0.93)	2.30 (.130)	24.81 (.000)
<i>IND</i> <sub><i>T</i></sub>	1.1 (1.88)	.01 (.89)	50.00 (4.34)	1.47 (.220)	0.79 (.375)
<i>MKT</i> <sub><i>T</i></sub>	0.35 (0.53)	.056 (2.83)	5.91 (0.44)	0.35 (.550)	7.99 (.006)
<i>VAR</i> <sub><i>T</i></sub>	1.67 (1.15)	0.17 (3.41)	21.32 (1.60)	0.33 (.560)	11.66 (.001)
<i>CORR</i> <sub><i>T</i></sub>	20.50 (4.98)	-0.04 (0.60)	16.98 (1.30)	2.37 (.120)	0.37 (.540)

*Notes.* This tables reports estimates of the parameters of the model of the variance and correlation series with a deterministic time trend. All the variables are defined as in the text. All the series are semi-annual (annualised). *DW* denotes the Durbin-Watson statistics of the static model. The point estimates of the  $\alpha$ ,  $\delta$ , and  $\beta$  parameters are multiplied by 100 to improve legibility. The rightmost columns report the Durbin's *h*-statistic of the null that the residuals are not first-order autocorrelated and the Wald statistic (in both cases with the associated significance levels) of the restriction that  $\delta$  is equal to zero. All the Wald and t-test statistics, standard errors and significance levels have been computed using a Newy-West adjusted variance-covariance matrix with Parzen weights to correct for error heteroskedasticity and autocorrelation. The sample period is 1974-2004. The estimated model is the following ( $u_T$  denotes an error term):

$$y_T = \alpha + \delta T + \beta y_{T-1} + u_T$$

### Short Run Dynamics

There is a potentially rich set of short run dynamic interactions between the variance components. Following the general-to-specific methodology (see, *inter alia*, Mizon (1995) and Kearney (2000)) I first specify a vector autoregression (VAR) model of the relation between overall market variance,  $MKT_t$ , industry variance,  $IND_t$ , and idiosyncratic firm variance,  $FIRM_t$ .

$$A(L)y_T = u_T \quad (5.15)$$

with

$$A(L) = I_3 - A_1L - A_2L^2 - \dots - A_QL^Q,$$

$$E(u_T) = 0, \quad E(u_T u_T') = \Sigma, \quad E(u_T u_S') = 0, \text{ for } T \neq S, \quad E(y_T u_T) = 0$$

and

$$y_T = [MKT_T \quad IND_T \quad FIRM_T]'$$

This is a reduced form VAR representation in which  $y_T$  is the vector of variables,  $I_3$  is a  $(3 \times 3)$  identity matrix,  $A_q$  are  $(3 \times 3)$  coefficient matrices,  $u_T$  is a  $(3 \times 1)$  vector of white noise disturbance terms, and  $L^q$  denotes the lag operator (for example,  $L^q y_T = y_{T-q}$ ). This model allows us to examine the full range of interaction between the variables in the  $y_T$  vector, i.e. the overall market variance,  $MKT_T$ , the industry variance,  $IND_T$ , and the idiosyncratic firm variance,  $FIRM_T$ . A convenient feature of the VAR representation in (5.15) is that it can be estimated by ordinary least squares, which yields consistent and asymptotically efficient estimates of the  $A_q$  matrices because the right-hand-side variables are predetermined and are the same in each equation of the model.

The first step in the estimation process is to decide on the appropriate lag length ( $Q$ ). The Akaike Information Criterion (AIC) suggests the inclusion of 3 lags, and the Swartz Bayesian Criterion (SBC) suggests 1 lag. Since a Likelihood-Ratio test



(LR) indicates that increasing the lag length from 1 to 3 produces a significant improvement in the overall model fit, I include 3 lags of each variable<sup>53</sup> ( $Q = 3$ ). This lag length selection tests are reported in Panel A of Table 5.4. I next perform block-exogeneity tests on the  $MKT_T$ ,  $IND_T$  and  $FIRM_T$  series to determine whether lags of one variable Granger-cause any of the others. If all lags of one variable can be excluded from the equations of the other two variables, we can model the latter using a 2-variable VAR. I test these restrictions using a likelihood ratio (LR) statistic, modified by Sims's (1980) multiplier correction to improve the small sample properties of the test. This test statistic is distributed as chi-squared with degrees of freedom equal to the number of lags excluded from each equation in the restricted system. Panel B of Table 5.4 presents the results. The only block exogenous variable is  $FIRM_T$ <sup>54</sup>. Moreover, from Panel C of the Table,  $MKT_T$  Granger-causes  $IND_T$  whereas the latter Granger-causes both  $MKT_T$  and  $FIRM_T$ .

Since the lags of both  $MKT_T$  and  $IND_T$  cannot be excluded from the equations of the other two variables, we must model the system as a trivariate VAR. To identify the structural model from the estimated reduced form, I impose the restrictions that  $IND_T$  does not have contemporaneous effects on  $MKT_T$ , and  $FIRM_T$  does not have contemporaneous effects on  $MKT_T$  and  $IND_T$ . The structural model is therefore written as follows.

$$B \begin{bmatrix} MKT_T \\ IND_T \\ FIRM_T \end{bmatrix} = \Gamma_0 + \sum_{q=1}^3 \Gamma_q L^q \begin{bmatrix} MKT_T \\ IND_T \\ FIRM_T \end{bmatrix} + \begin{bmatrix} \varepsilon_{MKT,T} \\ \varepsilon_{IND,T} \\ \varepsilon_{FIRM,T} \end{bmatrix} \quad (5.16)$$

Here,  $\Gamma_0$  is a vector of constants and  $B$  and  $\Gamma_q$  are, respectively, the contemporaneous and lagged structural coefficient matrices. The elements along the main diagonal of  $B$  are equal to 1. Denoting by  $L$  a Cholesky lower triangular matrix such that

<sup>53</sup> Moreover, a Likelihood Ratio test does not reject the restriction that the lag length is one instead of two (the Chi-squared statistics is 7.83 with significance level 0.550).

<sup>54</sup> The significance level of  $MKT_T$  is only slightly higher than the 5 percent level.

$\Sigma = LL'$  and by  $D$  the diagonal matrix of the structural errors standard deviations (the elements along the main diagonal of  $L$ ), I impose the restriction that  $DD = B^{-1}\Sigma(B^{-1})' \equiv B^{-1}LL'(B^{-1})'$ . It follows that  $D = B^{-1}L$  and  $B = LD^{-1}$ . With estimates of  $\Sigma$  in hand from the residuals of the reduced form model in (5.15), I solve for  $B$  and obtain the following point estimates of its elements:

$$B = \begin{bmatrix} 1 & 0 & 0 \\ -.451 & 1 & 0 \\ -.436 & -.119 & 1 \end{bmatrix} \quad (5.17)$$

The elements of  $B$  suggest that there is a positive contemporaneous influence of  $MKT_T$  on both  $IND_T$  and  $FIRM_T$ , and of  $IND_T$  on  $FIRM_T$ . Table 5.5 reports the corresponding variance decomposition of the Euro area market variance and average industry and firm level variance series. A large portion of the variance of  $IND_T$  and  $FIRM_T$ , over 30 percent one period ahead, is explained by variation in  $MKT_T$ , whereas only 5.7 percent of the latter is explained by variation in  $IND_T$  after 3 periods and none by  $FIRM_T$ . Interestingly, the impulse-response functions, reported in Figure 5.2, show that the most intense interactions take place with a three period lag (corresponding to 18 months).

Comparing my findings to those reported by CLMX (2001), it emerges that both systematic and industry level variance play a more important role in Europe than in the United States whereas firm volatility is more important in the United States. This suggests that their role in forecasting exercises, which might be relevant in pricing applications and asset allocation decisions as suggested by, *inter alia*, Goyal and Santa Clara's (2003) work, is different depending on whether the stocks are drawn from a European rather than a United States sample.



**Table 5.4**  
**Short Run Volatility Components Dynamics**  
**Reduced Form Model**

*Panel A*  
(Lag-length Selection)

Lags	AIC	SBC	LR	p-value
1	-23.702	-23.243*		
2	-23.558	-22.755	10.785	0.290
3	-23.900*	-22.753	35.122	0.000
4	-23.777	22.286	11.861	0.221
5	-23.674	-21.839	12.852	0.169
6	-23.586	-21.406	13.591	0.137

*Panel B*  
(Block-exogeneity Tests)

Variable	$\ln \Sigma_{UR} $	$\ln \Sigma_R $	Chi-Squ.(6)	Sig.
MKT <sub>T</sub>	-16.751	-16.525	11.730	0.068
IND <sub>T</sub>	-16.693	-16.350	17.847	0.006
FIRM <sub>T</sub>	-16.458	-16.315	7.431	0.282

*Panel C*  
(Granger Causality Tests)

Dep. Variable	Lags	F-Statistic	Sig.
MKT <sub>T</sub>	MKT <sub>T-q</sub>	1.713	0.176
	IND <sub>T-q</sub>	7.176	0.000
	FIRM <sub>T-q</sub>	0.130	0.941
IND <sub>T</sub>	MKT <sub>T-q</sub>	2.822	0.048
	IND <sub>T-q</sub>	8.812	0.000
	FIRM <sub>T-q</sub>	1.596	0.202
FIRM <sub>T</sub>	MKT <sub>T-q</sub>	0.120	0.947
	IND <sub>T-q</sub>	3.947	0.013
	FIRM <sub>T-q</sub>	1.188	0.324

*Notes.* Panel A of this table reports, for the trivariate VAR system of MKT<sub>T</sub>, IND<sub>T</sub> and FIRM<sub>T</sub> the AIC, the SBC and the Likelihood Ratio (LR) test statistics. The latter is constructed as the change in the likelihood function each time the lag length is incremented. The p-value refers to the LR statistic. Panel B reports the log-determinants of the unrestricted ( $\ln|\Sigma_{UR}|$ ) and restricted ( $\ln|\Sigma_R|$ ) VAR systems where the variable specified in the left-most column is restricted to be block-exogenous. The Chi-Squared statistic is computed as  $(T - c)(\ln|\Sigma_R| - \ln|\Sigma_{UR}|)$ , where  $T = 61$  and  $c$  is Sims' (1980) multiplier correction. Panel C reports Granger-causality tests of the null that all the lags of a variable can be excluded from the equation of the dependent variable. All the variables are linearly de-trended. The sample period is 1974-2004.

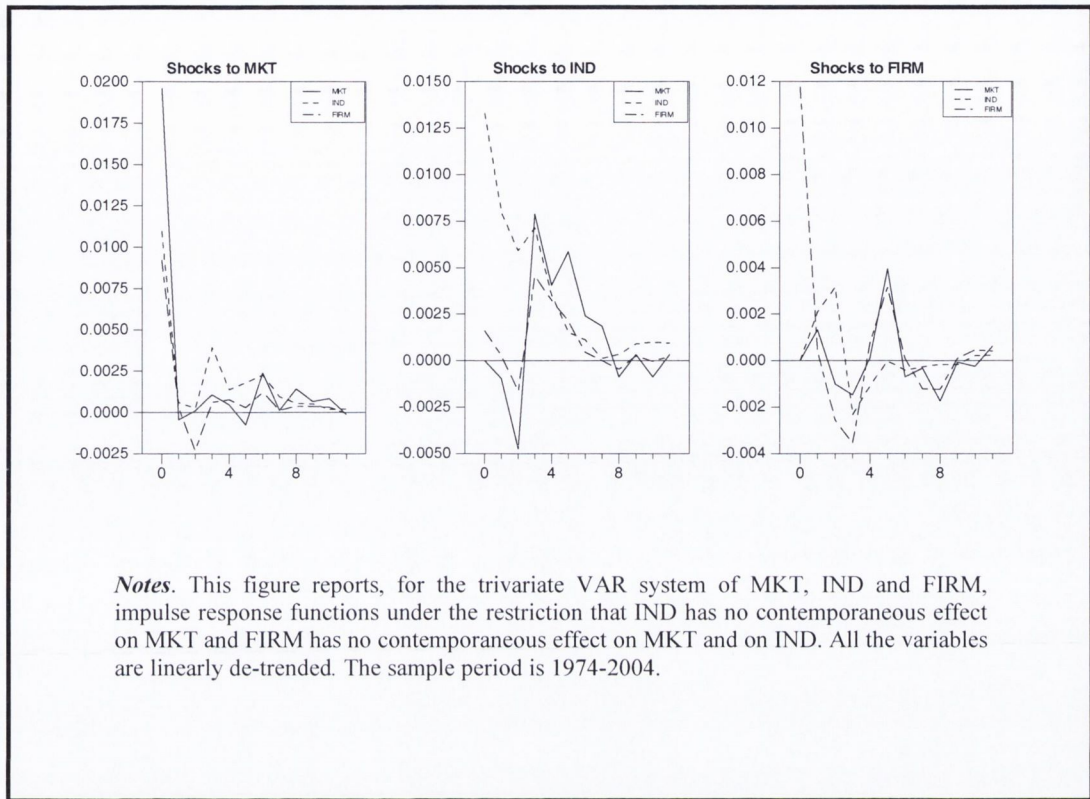
**Table 5.5**  
**Short Run Volatility Components Dynamics**  
**Structural Model**

Series	St. Error	Step	MKT <sub>T</sub>	IND <sub>T</sub>	FIRM <sub>T</sub>
MKT <sub>T</sub>	1.95	1	100.0	0.0	0.0
		2	99.3	0.2	0.4
		3	93.6	5.7	0.6
IND <sub>T</sub>	1.96	1	40.3	59.6	0.0
		2	32.9	65.8	1.1
		3	29.3	67.1	3.4
FIRM <sub>T</sub>	2.02	1	37.5	1.1	61.2
		2	37.5	1.1	61.2
		3	37.5	2.2	60.2

*Notes.* This table reports, for the trivariate VAR system of MKT<sub>T</sub>, IND<sub>T</sub> and FIRM<sub>T</sub>, the percentage of the variance of the series reported in the first column explained by the series reported at the top of each row. The variance decomposition imposes the restriction that IND<sub>T</sub> has no contemporaneous effect on MKT<sub>T</sub> and FIRM<sub>T</sub> has no contemporaneous effect on MKT<sub>T</sub> and on IND<sub>T</sub>. All the variables are linearly de-trended. The sample period is 1974-2004.



**Figure 5.2**  
**Short Run Volatility Components Dynamics**  
**Structural Model**



## 5.5. US vs. EMU Series

In order to examine more closely the extent to which my constructed EMU variance and correlation series share similar features to those displayed by the series used by CLMX (2001) for US markets, I construct comparable albeit shorter variance and correlation series from CLMX (2001) data<sup>55</sup>. To do so, I simply aggregate at a semi-annual frequency the monthly CLMX (2001) market, industry and firm-level variance series constructed from weekly returns from 1974 to 1997 and I then multiply the results by a factor of two to annualise<sup>56</sup>. The average idiosyncratic variance series is computed, according to (5.11), as the sum of the average industry and firm-level variance series. Since average total stock variance  $VAR_t$  and the square of average total volatility  $VOL^2$  are likely very similar, I use the former instead of the latter in (5.12) to construct an approximate average stock correlation series, not available in the CLMX (2001) study. I obtain 48 non overlapping semi-annual value-weighted variance and correlation data points ( $T = 1, 2, \dots, 48$ ) computed from weekly returns data. Panel B of Figure 5.1 reports the ratio of firm to industry variance and Panel C plots the average stock correlation. A striking difference between the US and EMU series is the relative importance of industry and firm level volatility over time. In the US, firm-level volatility becomes the largest component of idiosyncratic volatility much earlier than in the EMU markets. In the latter, industry level volatility is still the largest component of idiosyncratic volatility for much of the 70s and 80s, as shown by Panel B of Figure 5.1.

## 5.6. What Might Explain Volatility and $R^2$ Trends?

CLMX (2001) and Wei and Zhang (2003), among others, suggest a number of circumstances that could explain the rise of idiosyncratic volatilities. The first obvious explanation is the tendency of conglomerates to break up into more specialized businesses, interpreted as a shift from internal to external capital

---

<sup>55</sup> I thank CLMX (2001) for kindly making their constructed variance series available.

<sup>56</sup> Notice, however, that  $MKT_T$  constructed from CLMX (2001) data is the variance of the market portfolio excess-return.



markets. Dennis and Strickland (2005) provide direct empirical evidence that increasing firm focus on specialized business is a significant determinant of the secular rise in firm-level volatility. The tendency to issue stocks at an earlier stage of the company life-cycle and Changes in executive compensation schemes that result in cash-flow volatility could also contribute to explain this phenomenon. These explanations could also account for why most of the increase has occurred in firm level rather than industry level volatilities. The argument that firm-level volatilities have increased because of the tendency towards less diversified conglomerates, however, applies less well to Euro area than to US stocks because it also implies a decrease in average correlations. Leverage is also an unlikely candidate to explain the rise in stock volatilities because, as a result of a secular tendency towards the disintermediation of financial transactions, it has declined over time both in the US and in the Euro area. Dennis and Strickland (2005), noting that there is an increase in idiosyncratic volatility following both positive and negative returns, also ruled out leverage as a possible cause for the increase in idiosyncratic volatility in US data. Brown and Kapadia (2005) suggest that the increase of idiosyncratic volatility can be explained by new listing by riskier companies or, more specifically, by a riskier sub-sample of the economy becoming publicly traded. They find that this phenomenon accounts for most of the rise of idiosyncratic volatility in the US. Examining this possibility using Euro area data is a useful extension of the present work that I leave for future research.

Under a more behavioural perspective, divergence between institutional and individual investors' sentiment, coupled with the increasing institutionalization of equity ownership and possible herding of mutual fund managers, could explain more trading and more volatile individual stock prices. For example, Xu and Malkiel (2003) and Dennis and Strickland (2005) find evidence of a positive relation between US idiosyncratic volatility and institutionalization of the ownership of US stocks. Morck, Yeung and Yu (2000) and, more recently, Li and Myers (2005), suggest a negative relation between the explanatory power of the market model and factors such as the degree of investor protection and the transparency of the agency



relationships between insider managers and outsider investors. From this perspective, the finding of a low average correlation and hence of a low market  $R^2$  is consistent with the generally good level of investor protection and transparency in Euro area stock markets.

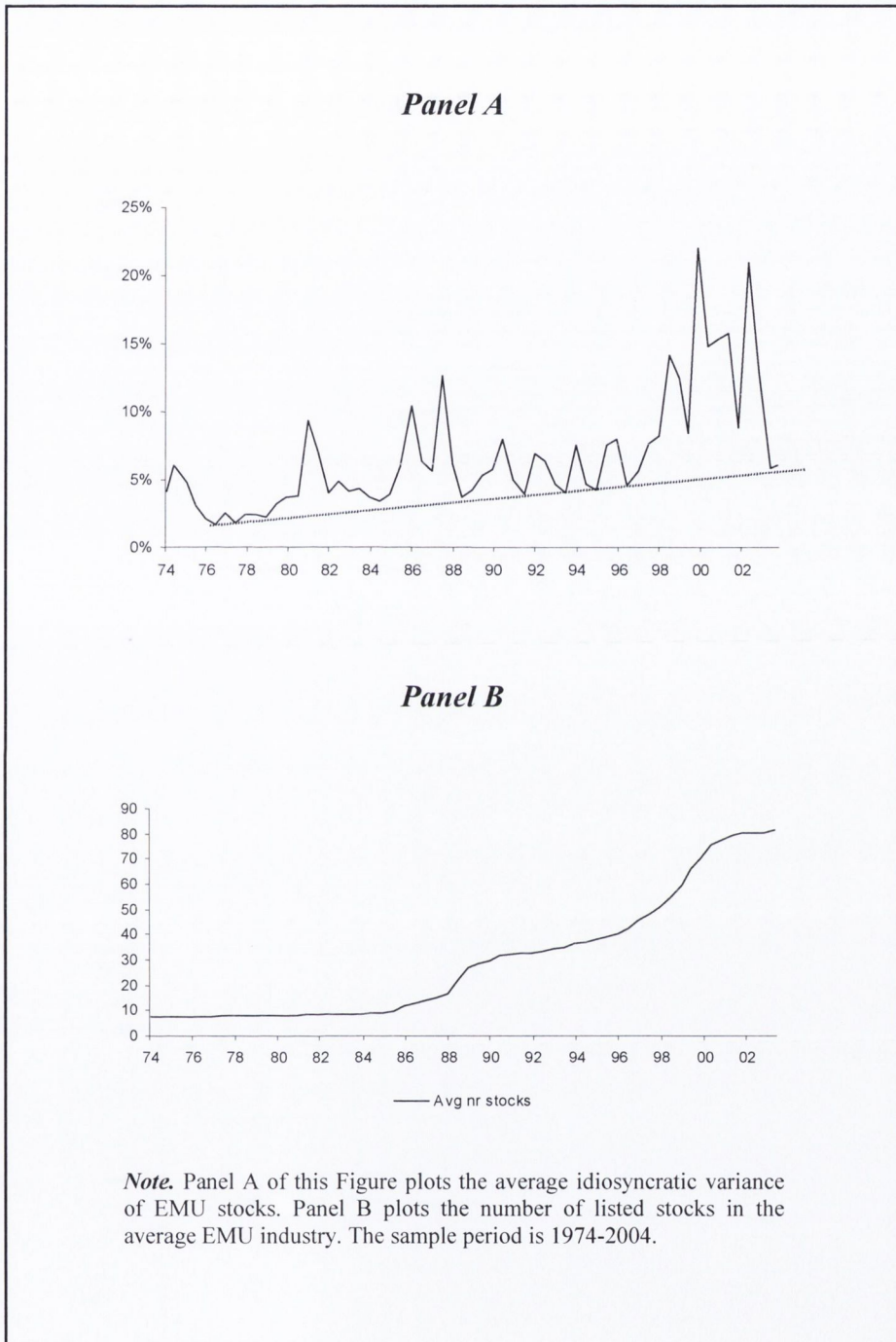
A further possibility is that the rise of idiosyncratic volatility from the end of the 1990s to the first years of the present decade might be a one-off episode rather than the result of a long-run trend. For example, a recent study by Brandt, Brav and Graham (2005) argues that the rise of idiosyncratic volatility in the United States during the same period is related to a speculative episode and that it can be explained on the basis of excess-trading by individual investors. Visual inspection of Panel A of Figure 5.3, however, suggests that, while idiosyncratic volatility reverted in the second semester of 2003 and the first semester of 2004 to pre-1997 lows, it did increase steadily over the sample period. More generally, the spike in idiosyncratic volatility towards the end of the sample period rises the possibility of a regime switch. In principle, this behaviour could be modelled within a switching ARCH/GARCH specification, allowing for different volatility regimes as in Edwards and Susmel (2000) and a time trend. This would yield efficient estimates of the time trend coefficient but, as it would require the simultaneous estimation of the mean and variance equations for all the 3515 stocks in my sample, it would be computationally unfeasible. A possible, albeit less efficient, alternative would be to allow for a regime switch directly in the mean of the unconditional volatility estimates used in this Chapter, e.g. the estimates based on (5.14). I leave this interesting development for future research.

Turning to the components of idiosyncratic volatility, the reason why industry level volatility is much larger than idiosyncratic volatility for much of the 70s and 80s, as shown by Panel B of Figure 5.1, is the limited cross-sectional dispersion within industries due to the small number of listed stocks. Unlike in the more mature United States markets, European industry indices initially comprised a small number of stocks with quite similar firms. As shown in Panel B of Figure 5.3, in 1974 the



number of stocks in the average industry index was less than 10, rising to about 30 by the end of the 1980s and since then it has grown steadily. In 2004 there were about 80 stocks in the average Euro area industry index. This also explains why, as reported in Panel C of Figure 5.1, the average correlation amongst Euro area industries is initially very similar to the average correlation amongst Euro area stocks, but the former increased relative to the latter from the mid 1980s.

**Figure 5.3**  
**Diversification in Industry Indices**





## 5.7. Second vs. Other Moments

Having studied the relation between the market and idiosyncratic variance components, I now focus on the relation between these series, aggregate correlation and market returns. In particular, I examine how their dynamics explain the asymmetry in the distribution of aggregate returns and the nature and strength of the risk-return trade-off. To this end, I employ again the VAR methodology. I first search for reduced form VAR models that best capture the interaction between the series and then I identify the underlying structural relations imposing simple restrictions on the structure of the residuals variance-covariance matrix dictated by economic theory and intuition. The reduced form models are systems of equations of the following form:

$$y_T = A_0 + \sum_{q=1}^Q A_q L^q y_T + u_T \quad (5.18)$$

Here,  $y_T$  denotes the vector of variables under consideration,  $A_0$  is a vector of constants,  $A_q$  are coefficient matrices and  $u_T$  is a vector of error terms. The structural form models are defined as follows:

$$B y_T = \Gamma_0 + \sum_{q=1}^Q \Gamma_q L^q y_T + \varepsilon_T \quad (5.19)$$

Here,  $B$  and  $\Gamma_q$  are structural coefficient matrices,  $\Gamma_0$  is a vector of constants,  $\varepsilon_T$  denotes the vector of structural errors, assumed to be uncorrelated with each other on a contemporaneous basis. All the vectors and matrices in (5.18) and (5.19) are conformable according to the rules of matrix algebra.

### *Aggregate Second Moments and Systematic Skewness*

To capture how the asymmetry of the multivariate distribution of Euro area stock returns arises from the dynamic interaction of aggregate first and second moments, I specify  $y_T$  as follows:

$$y_T = [CORR_T \quad VAR_T \quad R_{m,T}] \quad (5.20)$$

The system of equations defined by (5.18) and (5.20) is a trivariate VAR reduced form model of average stock correlation, total stock variance and the market return. Both the AIC and SBC, reported in Panel A of Table 5.6, suggest to include only one lag. Hence, I set  $Q = 1$ . The reduced form coefficient estimates  $A_1$ , the impact multipliers, are reported in Panel B of the Table.

To identify the contemporaneous and lagged relations between the variables, I impose a triangular Cholesky decomposition of the variance-covariance matrix of the reduced form residuals and I experiment with various orderings of the variables. In Panel C of Table 5.6 I report the structural coefficients for the orderings  $R_{mT} \rightarrow CORR_T \rightarrow VAR_T$ ,  $R_{mT} \rightarrow VAR_T \rightarrow CORR_T$  and  $CORR_T \rightarrow VAR_T \rightarrow R_{mT}$ . The latter is the only one not to imply the implausible prediction that volatilities and correlations move systematically in opposite directions. It is also consistent with the circumstance that, in the estimated reduced form model given by (5.18) and (5.20), the coefficient of determination  $R^2$  of the  $R_{mT}$  and of the  $VAR_T$  equations is 8.48 and 8.08 percent (3.58 and 3.15 percent adjusting for the degrees of freedom) respectively, whereas it is just 3.47 percent (almost zero adjusted) for  $CORR_T$ , suggesting that variation in average correlation drives variation in the market return and in the volatility of the typical stock and not the other way around. The structural coefficients suggest a rich contemporaneous interaction between the variables, whereas the only noticeable lagged interaction, computed solving  $A_1 = B^{-1}\Gamma_1$  for  $\Gamma_1$ , is the effect of average correlation on future returns.



To better gauge the relative importance of the lagged and contemporaneous relations between first and second moments, I regress in Table 5.7 my constructed variance and correlation series on contemporaneous and lagged values of the market return. At this (relatively low) frequency, there is no evidence that past negative returns lead to higher volatility for the average stock or to stronger average correlation among the stocks as implied, at lower frequencies (daily or weekly), by models of asymmetric conditional volatility and correlations such as the DCC-GARCH model employed by Cappiello, Engels and Sheppard (2003). Thus, there is no direct evidence of 'contagion' across stocks in the Euro area. This does not mean, however, that phenomena of contagion, as defined for example by King and Wadhvani (1990), do not take place in the Euro area. These phenomena, while widespread and possibly intense, might well be short lived and therefore they might be difficult to detect in semi-annual data. There is, however, clear evidence that higher volatilities are associated on average with low values of market returns, and that in these circumstances, average correlation also tends to be high (negative contemporaneous relationship of, respectively, average variance and correlation with market returns). On the basis of (5.12), this explains why market variance tends to be high when market returns are low (negative contemporaneous relationship of market variance with market returns). It therefore explains why the distribution of market returns is skewed to the left<sup>57</sup>.

---

<sup>57</sup> The skewness of the distribution of total weekly returns on the EMU Datastream equity index over the period 1974-1997 and 1974-2004 reported by Kearney and Poti (2003) is, respectively, -0.47 and -0.64 (significantly different from zero at, respectively, the 7% and the 13% level).

**Table 5.6**  
**Second Moments vs. Market Return**  
**(VAR)**

*Panel A*  
(Lag-length Selection)

Lags	AIC	SBC
1	-15.1832*	-14.7244*
2	-15.0096	-14.2066
3	-14.9957	-13.8485
4	-15.0665	-13.5752

*Panel B*  
(Reduced Form Coefficients)

$$A_1 = \begin{vmatrix} 0.196 & 0.358* & -0.208 \\ 0.040 & 0.167 & 0.112 \\ 0.013 & -0.116 & 0.322* \end{vmatrix}$$

*Panel C*  
(Structural Coefficients)

Ordering  $R_{mT} \rightarrow CORR_T \rightarrow VAR_T$ :

$$B = \begin{vmatrix} 1 & 0 & 0 \\ 0.310 & 1 & 0 \\ 0.149 & 0.179 & 1 \end{vmatrix} \quad \Gamma_1 = \begin{vmatrix} 0.196 & 0.358 & -0.208 \\ 0.101 & 0.278 & 0.047 \\ 0.035 & -0.093 & 0.270 \end{vmatrix}$$

Ordering  $R_{mT} \rightarrow VAR_T \rightarrow CORR_T$ :

$$B = \begin{vmatrix} 1 & 0 & 0 \\ 0.205 & 1 & 0 \\ 0.175 & 0.655 & 1 \end{vmatrix} \quad \Gamma_1 = \begin{vmatrix} 0.196 & -0.208 & 0.358 \\ 0.081 & 0.069 & 0.241 \\ 0.020 & 0.211 & -0.164 \end{vmatrix}$$

Ordering  $CORR_T \rightarrow VAR_T \rightarrow R_{mT}$ :

$$B = \begin{vmatrix} 1 & 0 & 0 \\ -0.252 & 1 & 0 \\ 0.270 & 0.844 & 1 \end{vmatrix} \quad \Gamma_1 = \begin{vmatrix} 0.167 & 0.112 & 0.040 \\ -0.159 & 0.294 & 0.002 \\ 0.304 & 0.093 & 0.218 \end{vmatrix}$$

*Notes.* Panel A of this table reports, for the trivariate VAR system of  $R_{mT}$ ,  $CORR_T$  and  $VAR_T$ , the AIC and the SBC. The sample period is 1974-2004. Panel B and Panel C report, respectively, the point estimates of the coefficients of the reduced form VAR (the symbol \* denotes significance at the 5 percent level) and the structural coefficients corresponding to various orderings of the variables.



**Table 5.7**  
**Second Moments vs. Market Return**  
**(Univariate Regressions)**

$$y = \text{const.} + \beta x + \text{regr. residual}$$

<i>y</i>	<i>x</i>	Coeff. Estimates [t-stat.]	<i>Adj. R</i> <sup>2</sup>
MKT <sub>T</sub>	R <sub>m,T-1</sub>	-0.01 [-0.41]	-0.01
	R <sub>m,T</sub>	-0.10 [-2.88]	0.29
VAR <sub>T</sub>	R <sub>m,T-1</sub>	-0.03 [-0.56]	-0.01
	R <sub>m,T</sub>	-0.21 [-2.92]	0.23
IDIO <sub>T</sub>	R <sub>m,T-1</sub>	-0.02 [-0.69]	-0.01
	R <sub>m,T</sub>	-0.11 [-2.69]	0.15
CORR <sub>T</sub>	R <sub>m,T-1</sub>	-0.02 [-0.25]	-0.01
	R <sub>m,T</sub>	-0.25 [-2.49]	0.09

**Notes.** This table reports regressions of the variance and correlation series on contemporaneous and lagged market returns over the period 1974-2004. The reported t-statistics (in squared brackets) are adjusted for heteroskedasticity and auto-correlation and regressions always include a constant. All variables are de-trended and all regressions include a constant.

### ***Predictive Relations***

To study the risk-return relation, I specify  $y_T$  as follows:

$$y_T = [MKT_T \quad IDIO_T \quad R_{m,T}] \quad (5.21)$$

Equations (5.18) and (5.21) define a reduced form VAR model of market variance, idiosyncratic variance and market return, i.e. of  $MKT_T$ ,  $IDIO_T$  and  $R_{m,T}$ . Both the AIC and the SBC, reported in Panel A of Table 5.8, suggest the inclusion of only one lag. I therefore estimate the reduced form VAR in (5.18) and (5.21) with  $Q = 1$ . To identify the structural risk-return relation, I impose a set of restrictions on the estimated reduced form VAR in a manner that is consistent with the market model of stock returns. In particular, I use a Cholesky decomposition of the VAR variance-covariance matrix that rules out contemporaneous effects of  $IDIO_T$  on  $MKT_T$  and of the market return on both variance series (this corresponds to the ordering  $MKT_T \rightarrow IDIO_T \rightarrow R_{m,T}$ ). In Figure 5.4, I plot the impulse response functions to visualize the impact of shocks to  $MKT_T$  and  $IDIO_T$  implied by this set of restrictions.

A shock to  $MKT_T$  has a large contemporaneous positive effect on  $IDIO_T$  and an even larger but negative effect on  $R_{m,T}$ . While the effect on  $IDIO_T$  fades quickly away, the lagged effect on the market return is marginally significant and of opposite sign (positive). Higher market variance, therefore, initially causes a drop in prices, but this effect turns positive the following period. This is consistent with the findings of a positive relation between market risk and expected return reported by Turner, Startz and Nelson (1989) and Harvey (1989).

A shock to  $IDIO_T$  has no initial impact on  $MKT_T$  (because of the restriction implied by the ordering of the variables). An increase of idiosyncratic volatility,  $IDIO_T$ , keeping market variance,  $MKT_T$ , constant, implies by (5.9) an increase in average total variance,  $VAR_T$ , and average total volatility,  $VOL_T$ . Thus it implies by (5.12) a



drop in average correlation. It also has a marginally significant positive effect on  $R_{mT}$ . In the following period, there is a marginally significant positive effect on  $MKT_T$  and a substantial negative effect on  $R_{mT+1}$ . This set of interaction implies a positive relation between average correlation and expected returns. Since market volatility is, by (5.12), proportional to average correlation, this is consistent with a positive relation between a systematic risk and expected return.

In Panel B of Table 5.8, I report predictive regressions of the market return using a constant and lagged variance series as regressors. Both market variance and average idiosyncratic variance predict market returns, but the relation with lagged market variance is positive while the relation with lagged idiosyncratic variance is negative. This is consistent with the preceding VAR analysis and with the impulse response functions in Figure 5.4. The relation between market returns and average idiosyncratic variance, however, is statistically significant at a conventional level in the 1974-2004 period, but not in the 1974-1997 period. Interestingly, this also obtains in the United States markets as reported by Guo and Savickas (2003). The significance levels of the reported t-statistics are confirmed by a simple bootstrap experiment<sup>58</sup>.

---

<sup>58</sup> Details on this bootstrapping experiment and the relevant RATS code are available upon request. The latter will also be available soon on my website, [www.valeripotenti.com](http://www.valeripotenti.com). The bootstrapping methodology is due to Efron (1979).

**Table 5.8**  
**Predicting the Market Return**

*Panel A*  
(VAR of  $MKT_T$ ,  $IDIO_T$  and  $R_{mT}$  - Lag-length Selection)

Lags	AIC	SBC
1	-19.298*	-18.839*
2	-19.137	-18.334
3	-19.165	-18.018
4	-19.209	-17.718
5	-19.206	-17.370
6	-19.047	-16.867

*Panel B*  
(Market Return Predictive Regressions)

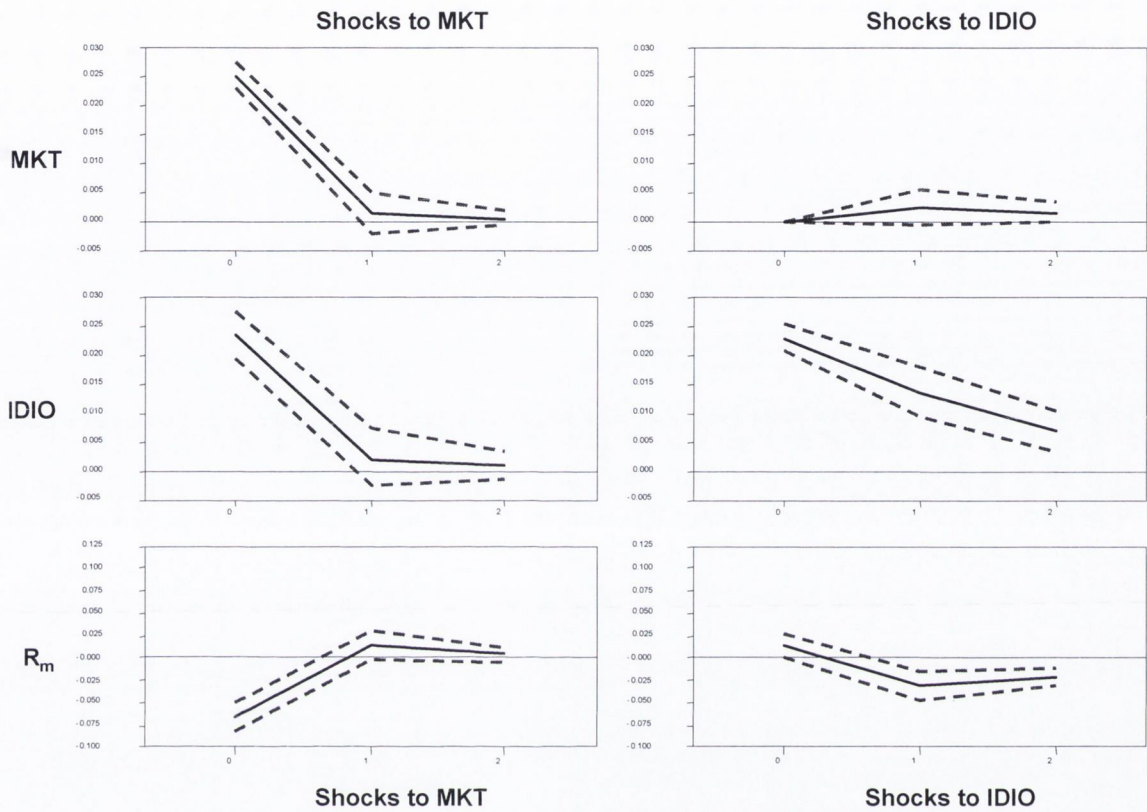
$$R_{mT} = const. + \beta_{MKT} MKT_{T-1} + \beta_{IDIO} IDIO_{T-1} + u_T$$

Restriction	$\beta_{MKT}$	$\beta_{IDIO}$	Adj. $R^2$
	<i>1974-1997</i>		
$\beta_{IDIO} = 0$	0.96 (1.97)		0.02
$\beta_{MKT} = 0$		-0.03 (-0.06)	0.00
	1.59 (2.02)	-0.78 (-1.15)	0.03
	<i>1974-2004</i>		
$\beta_{IDIO} = 0$	0.42 (0.62)		0.01
$\beta_{MKT} = 0$		-0.69 (-2.28)	0.04
	1.95 (2.67)	-1.55 (-3.83)	0.12

*Notes.* Panel A reports, for the trivariate VAR system of  $MKT_T$ ,  $IDIO_T$  and  $R_{mT}$ , the AIC and the SBC.  $MKT_T$ ,  $IDIO_T$  are linearly de-trended. Panel B reports coefficients estimates and coefficient of determination of predictive regressions of Euro area market returns. In brackets are t-statistics adjusted for heteroskedasticity and auto-correlation and regressions always include a constant. The sample period is 1974-2004



**Figure 5.4**  
**Impulse Responses of Variance Series and Market Return to Volatility Shocks**



*Notes.* This figure plots the impulse response functions of the  $MKT_T$ ,  $IDIO_T$  and  $R_{mT}$  series to shocks to  $MKT_T$  and  $IDIO_T$ . The variance series are linearly detrended. The model is estimated under the restriction that  $IDIO_T$  has no contemporaneous effect on  $MKT_T$  and  $R_T$  has no contemporaneous effect on  $MKT_T$  and  $IDIO_T$ . The symbols retain the usual meaning as in the text. The sample period is 1974-2004. The 95% confidence bands are constructed using a Montecarlo integration procedure.

## 5.8. Summary and Conclusions

In this Chapter I studied both systematic and idiosyncratic volatility in the stock markets of the Euro area. I employed a variance decomposition methodology similar to the approach proposed by CLMX (2001) and I derived the approximate relation between market volatility, idiosyncratic volatility and average correlation. I applied this analytical framework to construct market variance, idiosyncratic variance (and its industry and firm components) and correlation series. I also constructed, for comparative purposes, analogous US variance and correlation series from CLMX (2001) data. Like in most empirical papers, my approach has been mainly descriptive, based on the application of econometric methods to infer the salient features of my constructed variance and correlation time series.

Regarding long term trends, my main findings are that, *first*, the variance of both the average European stock and of the Euro area market portfolio has increased over time and that a large portion of this increase is explained by a long-run trend. European stocks, therefore, have indeed become more volatile. One consequence of the increase of average idiosyncratic risk is that it takes increasingly more stocks to fully capture the benefit of diversification. *Second*, value-weighted average stock correlation is relatively stable as it tends to mean revert quickly to a (roughly) 20 percent long run mean after a shock. *Third*, idiosyncratic volatility accounts for the main portion of the variance of the typical stock. The potential benefits to diversification strategies are, therefore, substantial. Regarding short run dynamics, EMU variance series are best forecast by market variance, whereas US variance series are best forecast by idiosyncratic variance. Market and average idiosyncratic variance, as already documented by Goyal and Santa Clara (2003) and by Guo (2003) using US data, predict market-wide returns. In the final Chapter, I will discuss the implications of my findings for portfolio management and financial theory.



## **Chapter 6: Correlations of European Equity Returns**

### **6.1. Introduction**

This Chapter is based on a paper forthcoming in the review *Research in International Business and Finance*, Kearney and Poti (2005a). I investigate the correlation trends and dynamics in the equity markets of the European Monetary Union (henceforth, EMU) to ascertain how the scope for country and industry level diversification has changed over the period 1993-2002 and to infer the dynamics that drive its change over time. This is of interest to portfolio managers who invest in European equity markets. It is also of interest to regulators because the level and common dynamics of correlations have implications for the stability of the financial system. Moreover, modelling correlation dynamics is important in understanding a crucial source of asymmetry in the multivariate distribution of stock returns. From this perspective, while Chapter 4 studied the asset pricing implication of distributional asymmetries under the 3M-CAPM, one of the aims of this Chapter is to quantify the asymmetry in the multivariate return distribution generated by equity correlation dynamics.

In particular, I study the correlation between EMU country and industry indices over various sample periods. For comparison, I also study the correlation amongst a sample of individual EMU stocks. To estimate conditional correlation dynamics, I use the recently developed DCC-MV GARCH model of Engle (2001) and Engle and Sheppard (2002). I specify this model to facilitate testing for non-stationarity, structural breaks and asymmetric dynamics in the correlation process. I also extend the DCC-MV GARCH model to include a deterministic time trend. My application of the DCC-MV GARCH model demonstrates how this specification effectively overcomes the dimensionality problems that often occur when modelling stock returns variance-covariance matrices.

In Section 6.2, I describe my dataset. In Section 6.3, I preliminarily model correlations in an unconditional setting and I test for the presence of either a

stochastic or a deterministic time trend. I then model conditional correlations using the DCC-MV GARCH model and its extensions. Section 6.4 summarizes the main findings and draws together the conclusions.

## 6.2. Data and Summary Statistics

All the data is expressed in Euro. The country level data is taken from Bloomberg and consists of daily returns from 1993 to 2002 on the 5 national stock market indices with the heaviest capitalisation in the euro-zone at the end of my sample period, ie, the *DAX* (Frankfurt Stock Exchange), the *CAC40* (Paris Stock Exchange), the *MIB30* (Milan Stock Exchange), the *AMX* (Amsterdam Stock Exchange) and the *IBEX* (Madrid Stock Exchange). These series start on 31 December 1991 (except for the *MIB30*, which starts a year later). The *Eurostoxx50* is the leading European stock market index, and the futures contract on this index is one of the most liquid in the world. It commenced on 31 December 1991 with a base value of 1000, and it comprises 50 stocks from the companies with the heaviest capitalisation in the euro-zone countries. From these, I select the 42 stocks with a continuous return history from 1993 to 2001 and I obtain from Bloomberg returns for the same time period. They are all traded in one of the 5 stock markets included in the country level sample. Because the stock indices are rebalanced to reflect the capitalization of the constituent stocks and because I only select the 42 stocks with a continuous return history over the entire sample period, there is a potential survivorship bias in both my market index and stock samples. The bias, however, is likely not very important as the sample period under consideration is not very long (intuitively, most of the stocks included in the indices and in the stock sample would survive anyway).

Table 6.1 provides the usual set of summary statistics for the returns on the 5 market indices and the individual stocks included in the *Eurostoxx50* index at the end of the 19 September 2001 reshuffle<sup>59</sup>. The summary statistics reported are returns sample means, variances, skewness, kurtosis, the *Jarque-Bera* statistics to test for normality.

---

<sup>59</sup> The excluded stocks are also listed in Table 1 and indicated by ‘\*’s.



Only the significance levels of the skewness statistics are reported to save space. The kurtosis and the *Jarque-Bera* statistics are statistically significant for all stocks in the sample. Thus, as expected, returns exhibit significant departures from normality in most cases. Noticeably, index returns always display negative skewness, whereas the sign of the latter is mixed across individual stocks. Table 6.2 reports the market sector (following the classification adopted by Bloomberg) of the individual stocks and highlights the stocks that are dropped from the sample due to the lack of a continuous return history over the entire sample period.

From *Estoxx Ltd* I take daily and weekly total return EMU market sectors indices from 1987 to 2004. From *Datastream International Ltd.*, I take weekly total return data from 1987 to 2004 on the 35 Level 4 fixed history EMU industry indices<sup>60</sup>. These represent at least 75% of the market capitalisation of the relevant industry. Sectors and industries are listed in Table 6.3. I also use Datastream 5 Year Government bond clean price indices for Germany, France, Spain, Italy and the Netherlands for the period 1987-2004.

---

<sup>60</sup> From Datastream I also take daily and weekly total return on the country indices used by Cappiello, Engle and Sheppard (2003) with a view to replicating their results.

**Table 6.1**  
**Summary Statistics for Stock and Market Index Returns**

	<i>Mean</i>	<i>Std. Dev.</i>	<i>Skew</i>	<i>Sig.</i>	<i>Kurt.</i>	<i>JB</i>
<i>Panel A: Market Indices</i>						
DAX	12.33	34.10	-0.44	0.000	3.72	1564
CAC40	10.37	19.75	-0.15	0.001	1.88	389
MIB30	13.66	23.56	-0.07	0.188	2.08	417
AEX	13.84	18.10	-0.39	0.000	4.38	2121
IBEX	12.23	20.43	-0.28	0.000	2.82	881
EUROSTOXX50	13.23	18.03	-0.29	0.000	3.65	1462
<i>Panel B: Individual Stocks</i>						
ABN AMRO	19.10	27.57	-0.17	0.001	4.47	2104
AEGON	32.39	28.33	0.20	0.001	4.19	1848
AHOLD	22.72	25.84	0.26	0.000	2.83	865
AIR LIQUIDE	13.28	27.75	0.24	0.000	2.14	485
ALCATEL	7.68	44.33	-0.97	0.000	17.27	30517
ALLIANZ	16.46	30.45	0.13	0.009	6.76	4398
AVENTIS	21.74	32.79	0.47	0.000	4.56	1957
N.A.	19.58	31.34	-0.12	0.013	3.04	938
BCO BILBAO VIZ. ARGENTARIA	26.41	30.21	0.10	0.040	6.88	4696
BASF	17.87	27.39	0.36	0.000	4.37	1885
BAYER	15.36	26.79	-0.28	0.000	7.21	5031
BAYER. HYPO & VEREINSBANK	12.25	33.02	0.35	0.000	5.31	2755
BNP	10.83	35.28	0.33	0.000	3.21	889
BCO SANTANDER CENTRAL HISP	20.74	32.21	-0.46	0.000	7.29	5346
CARREFOUR SUPERMARCHE	20.93	29.28	0.02	0.623	2.98	896
DAIMLERCHRYSLER	-7.40	34.46	-0.01	0.868	1.74	96
N.A.	6.93	26.12	0.06	0.205	3.38	1153
DEUTSCHE BANK R	12.36	30.98	0.20	0.000	6.62	4228
DEUTSCHE TELEKOM	12.67	46.80	0.30	0.000	1.43	125
E.ON	15.66	26.46	0.22	0.000	3.28	1051
ENDESA	19.88	25.79	0.07	0.141	2.36	553
ENEL	-6.00	28.02	-0.10	0.335	2.15	101
ENI	19.59	28.55	0.13	0.039	1.33	113
FORTIS B	22.06	26.22	0.10	0.038	3.64	1343
FRANCE TELECOM	19.12	52.42	0.63	0.000	3.33	537
ASSICURAZIONI GENERALI	14.11	26.36	0.17	0.001	2.11	462
ING GROEP	27.16	28.55	-0.48	0.000	8.22	7153
L'OREAL	26.45	32.67	0.10	0.054	1.85	350
N.A.	11.31	33.50	0.40	0.000	4.11	1771
MUENCHENER RUECKVER R	29.12	40.74	-1.72	0.000	31.38	59805
NOKIA	92.62	49.62	-0.08	0.105	5.12	2624
PHILIPS ELECTRONICS	36.53	42.34	-0.18	0.000	3.92	1615
PINAULT PRINTemps REDOUTE	25.22	31.22	0.04	0.456	3.03	923
REPSOL YPF	16.54	24.85	0.63	0.000	6.29	4088
ROYAL DUTCH PETROLEUM	16.09	23.35	0.09	0.075	2.79	815
RWE	12.60	27.43	0.48	0.000	5.17	2659
SAINT GOBAIN	30.13	32.67	0.18	0.000	1.95	397
SAN PAOLO IMI	12.53	33.73	0.34	0.000	2.21	524
SIEMENS	16.96	32.02	0.27	0.000	6.54	4407
N.A.	16.00	32.75	0.07	0.152	3.12	983
SOC GENERALE A	13.94	30.62	0.08	0.127	2.30	539
SUEZ	12.31	26.83	0.37	0.000	2.86	855
TELECOM ITALIA	30.41	35.53	-0.26	0.000	5.23	2791
TELEFONICA	26.81	31.70	0.08	0.091	1.77	314
TIM	33.65	37.28	0.23	0.000	0.76	51
TOTAL FINA ELF	17.61	30.21	-0.03	0.527	1.59	256
UNICREDITO ITALIANO	17.85	37.28	0.76	0.000	4.33	2121
UNILEVER NV	15.45	23.98	0.31	0.000	6.45	4382
N.A.	10.23	30.21	0.18	0.000	2.74	770
VOLKSWAGEN	15.34	31.90	0.07	0.161	3.86	1532

*Notes.* The table reports summary statistics for the five largest EMU stock market indices, for the Eurostoxx50 and for the stocks included in the latter on 23 November 2001. The sample period is 1993-2001. Mean and standard deviations are in percentage on a 1-year basis. *JB* denotes the *Jarque-Bera* statistics. The reported significance levels refer to skewness. The Kurtosis and the *JB* statistics are different from zero at the 0.1 percent level for all stocks in the sample.



**Table 6.2**  
**Stocks Included in the Eurostoxx50 Index**

	Company	Bloomberg Ticker	Market Sector	Weights
1	ABN AMRO	AABA NA	BAK	1.59%
2	AEGON	AGN NA	INN	1.55%
3	AHOLD	AHLN NA	NCG	1.87%
4	AIR LIQUIDE	AI FP	CHE	0.89%
5	ALCATEL	CGE FP	THE	1.02%
6	ALLIANZ	ALThe V GY	INN	2.49%
7	ASSICURAZIONI GENERALI	G IM	INN	2.15%
8	AVENTIS	AVE FP	HCA	3.48%
9	AXA UAP	N.A.	INN	2.00%
10	BASF	BAS GY	CHE	1.26%
11	BAYER	BAY GY	CHE	1.40%
12	BAYERISCHE HYPO & VEREINSBANK	HVM GY	BAK	0.75%
13	BCO BILBAO VIZCAYA ARGENTARIA	BBVA SM	BAK	2.39%
14	BCO SANTANDER CENTRAL HISP	SAN SM	BAK	2.46%
15	BNP*	BNP FP	BAK	2.37%
16	CARREFOUR SUPERMARCHE	CA FP	RET	1.97%
17	DAIMLERCHRYSLER*	DCX GY	ATO	1.86%
18	DEUTSCHE BANK R	DBK GY	BAK	2.13%
19	DEUTSCHE TELEKOM*	DTE GY	TEL	2.64%
20	E.ON	EOA GY	UTS	2.39%
21	ENDESA	ELE SM	UTS	1.14%
22	ENEL*	ENEL IM	UTS	0.83%
23	ENI*	ENI IM	ENG	2.22%
24	FORTIS B	FORB BB	FSV	0.98%
25	FRANCE TELECOM*	FTE FP	TEL	1.06%
26	GROUPE DANONE	N.A.	FOB	1.47%
27	ING GROEP	INGA NA	FSV	2.95%
28	L'OREAL	OR FP	NCG	1.52%
29	LVMH MOET HENNESSY	N.A.	CGS	0.55%
30	MUENCHENER RUECKVER R*	MUV2 GY	INN	1.70%
31	NOKIA	NOK1V FH	THE	5.63%
32	PHILIPS ELECTRONICS	PHIA NA	CGS	1.75%
33	PINAULT PRINTEMPS REDOUTE	PP FP	RET	0.49%
34	REPSOL YPF	REP SM	ENG	1.02%
35	ROYAL DUTCH PETROLEUM	RDA NA	ENG	7.63%
36	RWE	RWE GY	UTS	0.98%
37	SAINT GOBAIN	SAN FP	CNS	0.81%
38	SAN PAOLO IMI	SPI IM	BAK	0.70%
39	SANOFI SYNTHELABO	N.A.	HCA	1.81%
40	SIEMENS	SIE GY	THE	2.34%
41	SOC GENERALE A	SGO FP	BAK	1.46%
42	SUEZ	SZE FP	UTS	2.39%
43	TELECOM ITALIA	TI IM	TEL	1.19%
44	TELEFONICA	TEF SM	TEL	3.24%
45	TIM*	TIM IM	TEL	1.22%
46	TOTAL FINA ELF	FP FP	ENG	7.31%
47	UNICREDITO ITALIANO	UC IM	BAK	0.84%
48	UNILEVER NV	UNA NA	FOB	2.49%
49	VIVENDI UNIVERSAL	N.A.	MDI	3.07%
50	VOLKSWAGEN	VOW GY	ATO	0.54%

*Note.* This table reports the stocks included in the Eurostoxx50 and their capitalization weights as of the September 2001 reshuffle. Asterisks indicate that the series has been dropped from the sample. Descriptors for the market sectors are as follows (Stoxx's Industry Codes): BAK (Banks), ATO (Auto), INN (Insurance), TEL (Telecom), NCG ((Non-Cyclical Goods and Services), UTS (Utilities), CHE (Chemical), ENG (Energy), THE (Technology), FSV (Financials), HCA (Health Care), FOB (Food & Beverages), RET (Retailer), CGS (Cyclical Goods and Services), CNS (Construction), MDI (Media).

**Table 6.3**  
**Sector and Industry Indices**

---

Panel A	
( <i>estoxx</i> EMU Sectors)	
1	Basic Mater. (SXBSCR )
2	Cyclic. Goods (SXCYCR )
3	Non Cyclical Goods (SXNCYR )
4	Energy (SXENER )
5	Financial Services (SXFINR )
6	Health care (SEHCRR )
7	Industrial Goods (SXIDUR )
8	Technology (SXTECR )
9	Telecom (SETLSR )
10	Utility (SEUTIR )
11	Basic Mater. (SXBSCT )
12	Cyclic. Goods (SXCYCT )
13	Non Cyclical Goods (SXNCYT )
14	Energy (SXENET )
15	Financial Services (SXFINT )
16	Health care (SEHCRT )
17	Industrial Goods (SXIDUT )
18	Technology (SXTECT )
19	Telecom (SETLST )
20	Utilities (SEUTIT )

Panel B	
( <i>Level 4 Datastream</i> EMU Industries)	
1	Aerospace & Defence (AERSP)
2	Automobiles & Parts (AUTMB)
3	Banks (BANKS)
4	Beverages (BEVES)
5	Chemicals (CHMCL)
6	Construction & Building Materials (CNSBM)
7	Diversified Industrials (DIVIN)
8	Electricity (ELECT)
9	Electronic & Electrical Equipment (ELTNC)
10	Engineering & Machinery (ENGEN)
11	Food & Drug Retailers (FDRET)
12	Food Producers & Processors (FOODS)
13	Forestry & Paper (FSTPA)
14	Health (HLTHC)
15	Household Goods & Textiles (HHOLD)
16	Information Technology Hardware (INFOH)
17	Insurance (INSUR)
18	Investment Companies (INVSC)
19	Leisure & Hotels (LESUR)
20	Life Ass. (LIFEA)
21	Media & Entertainment (MEDIA)
22	Mining (MNING)
23	Oil & Gas (OILGS)
24	Personal Care & Household Products (PERSH)
25	Pharmaceuticals & Biotechnology (PHARM)
26	Real Estate (RLEST)
27	Retailers, General (RTAIL)
28	Software & Computer Services (SFTCS)
29	Speciality & Other Finance (SPFIN)
30	Steel & Other Metals (STLOM)
31	Support Services (SUPSV)
32	Telecom Services (TELCM)
33	Tobacco (TOBAC)
34	Transport (TRNSP)
35	Utilities, Other (UTILO)

---

*Notes.* The table reports the sector (Panel A) and industry (Panel B) indices used in this study.



### 6.3. EMU Country and Stock Correlations

I first construct unconditional estimates of EMU country and stock correlations based on (3.13). To do this, I compute the cross products of the standardised daily log-return  $R_{it}$  deviations from their sample means and sum them to obtain correlation measures for each pair of stocks  $i$  and  $j$  over non-overlapping monthly periods:

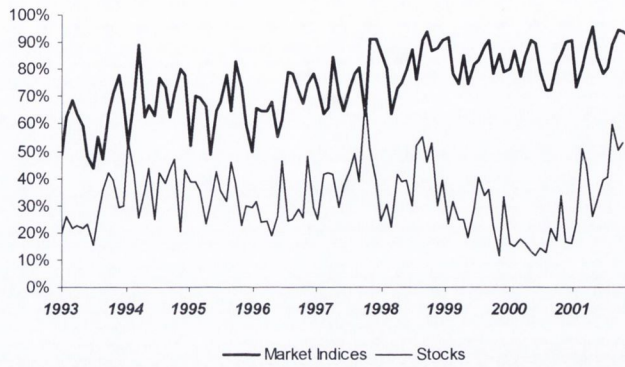
$$c_{i,j,T} = \frac{\sum_{k=1}^p (R_{i,t-k+1} - \bar{R}_{i,T})(R_{j,t-k+1} - \bar{R}_{j,T})}{\sqrt{\text{Var}(R_{i,t})_T \text{Var}(R_{j,t})_T}} / p \quad (6.1)$$

Here, using the convention that each month is made up of 21 trading days, I set  $p = 21$ . I then average correlations across market indices and stocks to construct their equally-weighted average correlation series.

$$CORR_T = \sum_{i=1}^n \sum_{j=1}^n \frac{1}{n^2} c_{i,j,T} \quad (6.2)$$

Here,  $n$  is either the number of market indices or of stocks, i.e.  $n = 5$  for the country-level market index sample and  $n = 42$  for the individual stock sample. The resulting average correlation series contains 105 monthly observations from 1993 to 2002 and are plotted in Figure 6.1. Thus,  $T = 1, 2, \dots, 105$ . The monthly average correlation amongst the indices shows a strong tendency to rise over time. The average stock correlation series instead does not display any obvious trend but rather it oscillates around a fairly stable long run mean of about 20 percent.

**Figure 6.1**  
**Average Market Index and Stock Correlations**



*Note.* This figure plots the unconditional estimates of the average correlation between the 5 largest stock market indices in the Euro area and the average correlation between 42 stocks included in the *Eurostoxx50* Index over the sample period 1993-2002.



### 6.3.1 Unit Root Tests

To test for the presence of a stochastic time trend, I conduct Dickey-Fuller (*DF*) and augmented Dickey-Fuller (*ADF*) tests allowing for up to 12 lags. As pointed out by Pesaran and Pesaran (1997), however, there is a size-power trade-off depending on the order of augmentation, and I consequently rely on the results provided by the tests performed at the lower orders of augmentation. The null of the *DF* and *ADF* test is  $H_0 : \rho = 1$ , with the estimate of  $\rho$  being obtained from the following equations:

$$y_T = \alpha_0 + \rho y_{T-1} + \sum_{i=1}^p \alpha_i \Delta y_{T-1} + u_T \quad (6.3)$$

$$y_T = \alpha_0 + \rho y_{T-1} + \sum_{i=1}^p \alpha_i \Delta y_{T-1} + \delta T + u_T \quad (6.4)$$

Here,  $y_T$  is the variable under consideration,  $p$  is the order of augmentation ( $p = 0$  for the *DF* and  $p \geq 1$  for the *ADF* test), the test statistic is  $\frac{(\hat{\rho} - 1)}{\sigma_{\hat{\rho}}}$ , where

$\hat{\rho}$  is the estimate of  $\rho$ ,  $\sigma_{\hat{\rho}}$  is its standard error and  $t$  is a time trend. Table 6.4 (Panel A) presents the results, reporting for brevity only the first 2 orders of augmentation. The *DF* and *ADF* tests reject the null of a unit root at the 5 percent level of significance in the average stock correlation series but not the null of a unit-root in the market index correlation series with 2 orders of augmentation and no deterministic time trend. However, using an *F*-test and the appropriate non-standard asymptotic distribution (Hamilton (1994)), I can reject at the 1 percent level the joint hypothesis that the deterministic time trend is equal to zero and the autocorrelation coefficient  $\rho$  is equal to unity. Moreover, by the Cauchy-Schwartz inequality, correlation is a bounded variable and therefore most likely stationary<sup>61</sup>. I therefore

---

<sup>61</sup> To formally check on this, one could repeat the unit-root tests using a transformation to unboundedness (Fisher transformation). Formally, this transformation is defined by  $z = \frac{1}{2} \ln\left(\frac{1+c}{1-c}\right)$ , where  $c$  is the correlation coefficient between two variables X and Y. If the distribution of these variables is bivariate normal and their correlation is  $\bar{c}$ , then  $z$  is approximately normally distributed

conclude that both correlation series are stationary, and in particular, that market index correlation is trend-stationary.

---

with mean  $\frac{1}{2} \ln\left(\frac{1+\bar{c}}{1-\bar{c}}\right)$  and standard deviation  $\frac{1}{\sqrt{N-3}}$ , where  $N$  is the sample size. The inverse of this transformation is  $z = \frac{e^{2z-1}}{e^{2z+1}}$ . I leave the examination of the Fisher transformation of the estimated correlation coefficients for future research.



**Table 6.4**  
**Unconditional Correlations**

Panel A					
(Unit Root Tests on Aggregate Correlations)					
	CV	DF	ADF1	ADF2	F-Test
<i>Country Indices</i>					
Intercept, no trend	-2.89	-4.95	-4.10	-2.67	620.01
Intercept and linear trend	-3.45	-7.62	-7.46	-5.57	(0.000)
<i>Individual Stocks</i>					
Intercept, no trend	-2.89	-5.68	-4.07	-3.30	
Intercept and linear trend	-3.45	-5.65	-4.04	-3.28	

Panel B					
(Specification and Wald-Type Tests)					
Static Model	Dynamic Model				
<i>DW-stat.</i>	$\alpha$ (%)	$\delta$ (%)	$\beta$	<i>h-stat.</i>	<i>Wald-stat.</i>
	( <i>t-stat.</i> )	( <i>t-stat.</i> )	( <i>t-stat.</i> )	( <i>sign.</i> )	( <i>sign.</i> )
<i>Country Indices</i>					
1.41	38.40	0.194	0.29	5.08	27.40
	(7.09)	(5.24)	(3.04)	(0.020)	(0.000)
<i>Individual Stocks</i>					
0.96	16.66	-0.008	0.51	2.60	0.05
	(4.07)	(0.22)	(5.95)	(0.100)	(0.820)

**Notes.** Panel A of this Tables reports Dickey-Fuller (*DF*) tests and augmented Dickey-Fuller (*ADF1* and *ADF2*, the numbers denoting the order of augmentation) tests for the presence of unit roots in the average country and stock unconditional correlations series. *CV* denotes the critical value at the 5 percent level. All variables are defined in the text. *F-test* denotes critical value and significance level (in brackets) of the test statistic under the null that the trend coefficient is zero and the series contains a unit root. Panel B reports estimates of the parameters of the model of the average country and stock correlations series with a deterministic time trend. *DW* denotes the Durbin-Watson statistics of the static model. All other columns report estimated coefficient and t-statistics for the dynamic model. The rightmost columns report the Durbin's *h*-statistic of the null that the dynamic model residuals are not first-order autocorrelated and the Wald statistic (in both cases with the associated significance levels) of the restriction that  $\delta$  is equal to zero. All the Wald-Test statistics, standard errors and significance levels have been computed using a Newy-West adjusted variance-covariance matrix with Parzen weights to correct for heteroskedasticity and autocorrelation. All variables are defined in the text. The sample period is 1993-2002.

*Static Model:*

$$y_T = \alpha + \delta T + u_T$$

*Dynamic Model:*

$$y_T = \alpha + \beta y_{T-1} + \delta T + u_T$$

### 6.3.2 Wald-Type Tests

To check on the possible presence of a deterministic time-trend, I conduct Wald-type tests of the restriction that the former does not help explain the variation in correlations. I first estimate a static model that includes among the regressors a deterministic time-trend coefficient but no lagged value of the dependent variable:

$$y_T = \alpha + \delta T + u_T \quad (6.5)$$

Since the Durbin-Watson (*DW*) statistic suggests that the residuals in (6.5) are auto-correlated, I estimate the following dynamic model to mitigate the error serial correlation problem:

$$y_T = \alpha + \beta y_{T-1} + \delta T + \varepsilon_T \quad (6.6)$$

I then conduct Wald-type tests of the restriction that the deterministic time trend coefficient is zero using Newy-West adjusted variance-covariance matrices to correct for heteroskedasticity and residual error autocorrelation. Table 6.4 (Panel B) presents the results. The time trend coefficient is large and significant only for average market index correlation. It explains an increase in the latter of about 2.5 percent per year. However, using Durbin's  $h^{62}$  statistic, I can only marginally (at the 10 percent level) reject the null that the residuals of average market index correlation from (6.5) are serially independent. The parameter estimates must therefore be treated with caution because, when lags of the dependent variable appear on the right hand side among the regressors, ordinary least squares (OLS) estimates are inconsistent in the presence of serially correlated errors.

---

<sup>62</sup> In the presence of lagged values of the dependent variables the *DW* test is biased toward acceptance of the null of no error auto-correlation. I therefore test for serial correlation of the error terms using Durbin's (1970) *h*-test. I use the generalised version of this test, developed by Godfrey and Breusch, based on a general Lagrange Multiplier test. Even though this procedure can detect higher order serial correlation, I only test the null of no first-order residual autocorrelation.



### 6.3.3 DCC-MVGARCH

Thus far I have applied an unconditional estimation methodology. This strategy has yielded useful insights but it has the main shortcoming that, while the average of squares and cross-products are consistent estimators of the second moments of the return distributions, they might be biased in small samples since they are ad hoc representations of the volatility and correlation processes. Moreover, the aggregation of daily data into lower frequency monthly data leads to a potentially accentuated small sample problem. It is, therefore, of considerable interest to apply the recently developed DCC-MVGARCH model of Engle (2001) and Engle and Sheppard (2002). This provides a useful way to describe the evolution over time of the second moments of large systems. In particular, I use the specification of the asymmetric DCC-MVGARCH (ADCC-MVGARCH) proposed by Cappiello, Engle and Sheppard (2003) in (3.42) and extend it to include a deterministic time trend:

$$C_t = \bar{C}(1 - \alpha - \beta) - \bar{S}\theta - \bar{t}(ii' - I)\delta_{trend} + \alpha\varepsilon_{t-1}\varepsilon'_{t-1} + \beta C_{t-1} + \theta S_{t-1} + \delta_{trend} t (ii' - I) \quad (6.7)$$

Here, the symbols retain their prior meaning,  $\bar{t}$  is the mid point of the sample period (the unconditional sample average of the values taken by the time trend variable),  $I$  is a conformable identity matrix and  $\delta_{trend}$  is the deterministic time-trend coefficient. To see why the inclusion of the deterministic time trend requires this modification of the correlation equation, consider for simplicity, but without loss of generality, the univariate case of a GARCH(1,1) with deterministic time trend,  $E_{t-1}(\varepsilon_t^2) = \alpha_0 + \alpha \varepsilon_{t-1}^2 + \beta E_{t-2}(\varepsilon_{t-1}^2) + \delta_{trend} t$ . Using the law of iterated expectations, its unconditional variance is:

$$E(\varepsilon_t^2) = \alpha_0 + (\alpha + \beta)E(\varepsilon_{t-1}^2) + \delta_{trend} E(t) = \alpha_0 + (\alpha + \beta)E(\varepsilon_t^2) + \delta_{trend} \bar{t} \quad (6.8)$$

Therefore, solving for the unconditional variance, we have that  $E(\varepsilon_t^2) = \frac{\alpha_0 + \delta_{trend} \bar{t}}{1 - \alpha - \beta}$

and  $\alpha_0 = [E(\varepsilon_t^2)(1 - \alpha - \beta) - \delta_{trend} \bar{t}]$ . The specification in (6.7) is a generalization to the multivariate case of this result.

Table 6.5 presents my ADCC-MVGARCH model estimates using daily data on the 5 market indices. I first estimate a simple symmetric specification of (6.7) with a deterministic time trend but no structural break. I label this specification Model 1. The estimated deterministic time trend coefficient turns out to be statistically significant but very small. Since it is economically negligible, I drop it from all subsequent specifications. I therefore estimate Model 2, which imposes on Model 1 the additional restriction that the time trend coefficient is zero.

Considering the apparent rise in average market index correlation that is visible in Figure 6.1, together with the lack of evidence of a significant deterministic time trend, I then test for the presence of either a stochastic trend (the correlation process is not stationary) or a structural break. To check the stationarity of the correlation process, I test the restriction that the news and persistence parameters  $\alpha$  and  $\beta$  sum to unity. The relevant LR test statistic and the associated significance level are reported at the bottom of Table 6.5 (Model 2 against Model 3). I reject the restriction that the parameters of the correlation process sum to unity and I conclude, therefore, that the correlation process is stationary.



**Table 6.5**  
**Conditional Country Correlation**

Panel A					
<i>Model</i>	<i>Restriction</i>	<i>Coefficient</i>	<i>Coefficient estimate</i>	<i>T-Ratio</i>	<i>p-value</i>
1	$Q_1 = Q_2$ $\theta = 0$	$Q_{1,2}$	.799		
		$\alpha$	.010	4.50	.000
		$\beta$	.982	180.97	.000
		$\delta_{Trend}$	.000	2.03	.041
2	$Q_1 = Q_2$ $\theta = 0$ $\delta_{Trend} = 0$	$Q_{1,2}$	.799		
		$\alpha$	.010	4.55	.000
		$\beta$	.985	223.82	.000
3	$Q_1 = Q_2$ $\theta = 0$ $\delta_{Trend} = 0$ $\alpha + \beta = 1$	$Q_{1,2}$	.799		
		$\alpha$	.007	12.72	.000
		$\beta$	.993	1807.09	.000
4	$\theta = 0$ $\delta_{Trend} = 0$	$Q_1$	.611		
		$Q_2$	.908		
		$\alpha$	.002	8.30	.000
		$\beta$	.970	589.68	.000
5	$\delta_{Trend} = 0$	$Q_1$	.611		
		$Q_2$	.908		
		$\alpha$	.002	3.90	.000
		$\beta$	.590	74.31	.000
		$\theta$	.090	7.69	.000

Panel B

<i>Unrestricted Model</i>	$\ln( \Sigma_{UR} )$	<i>Restricted Model</i>	$\ln( \Sigma_R )$	<i>LR Statistic</i>	<i>Significance Level</i>	<i>Restriction Rejection</i>
2	-5.0689	3	-5.0798	33.19	.000	Yes
4	-5.0665	2	-5.0689	25.15	.020	Yes
5	-5.0654	4	-5.0665	2.53	.112	No

$$LR = -T[\ln(|\Sigma_{UR}|) - \ln(|\Sigma_R|)] \sim \chi^2(I)$$

$T$  = number of observations (2,297)

$\Sigma_{UR}$  = covariance matrix of the residuals of the unrestricted model

$\Sigma_R$  = covariance matrix of the residuals of the restricted model

$\chi^2(I)$  = Chi-Squared distributions with  $I$  degree of freedom

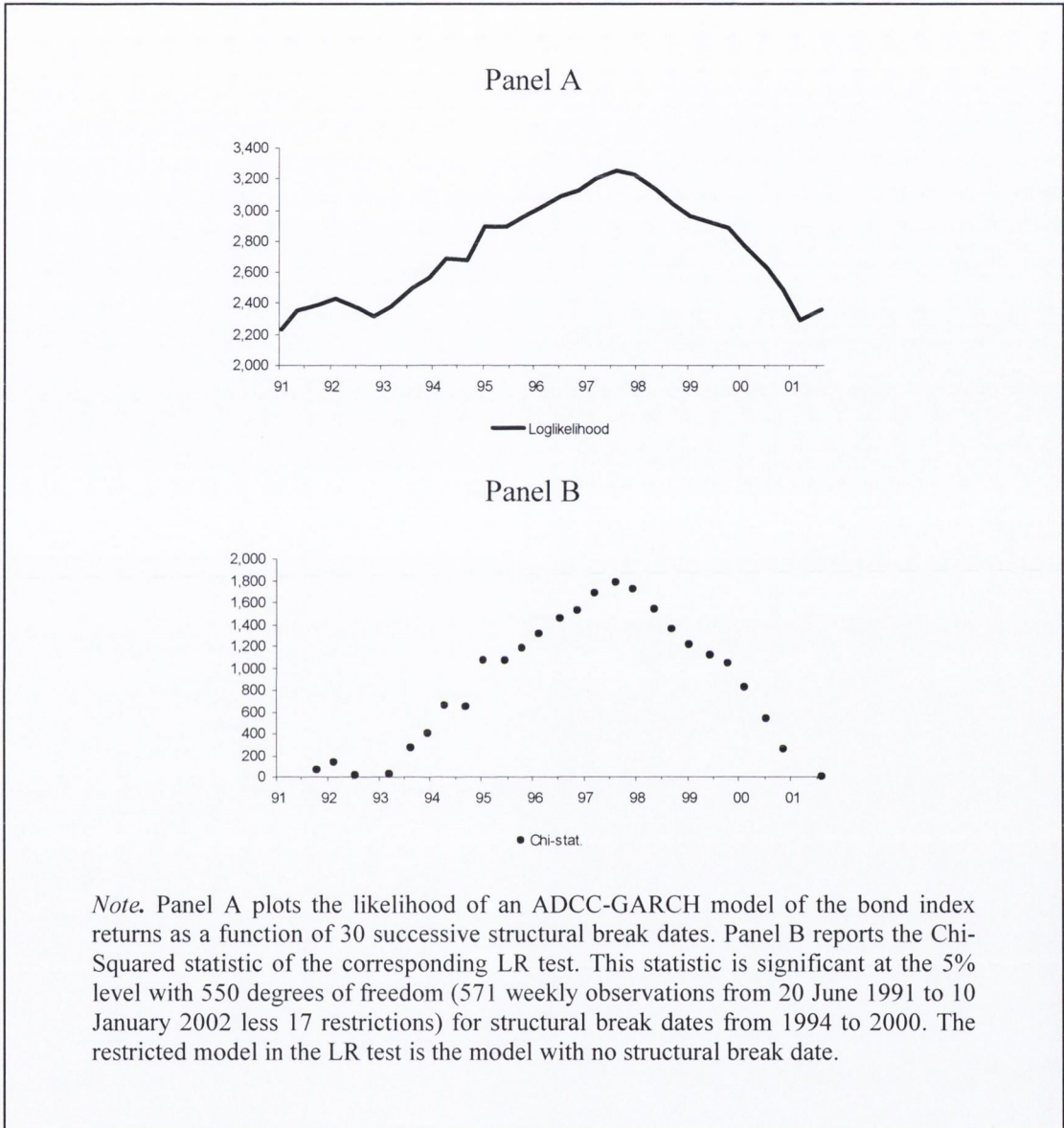
*Notes.* Panel A of this Table reports coefficients,  $t$ -statistics and  $p$ -values for various specifications of the ADCC-MVGARCH model of conditional correlations amongst the 5 largest Euro-zone market indices over the period 1993-2002. Panel B reports Likelihood Ratio (LR) test statistics and their significance level.

A structural break in the market index correlation process might, however, explain both the strong persistence of the series and its sharp increase over the sample period. A structural break in EMU interest rates correlations due to monetary policy convergence is a likely cause of a structural break in correlations at the stock market index level, as suggested for example by the study of Hardouvelis, Malliaropoulos and Priestley (2000) and Cappiello, Engle and Sheppard (2003). The plot of the likelihood of an ADCC-GARCH model of the Government bond index returns as a function of 30 successive structural break dates, as reported in Figure 6.2, peaks at the beginning (January) of 1998. The hypothesis that this might be the structural break date is intuitively appealing since it is roughly 12 months before the official introduction of the Euro and thus it accounts for the likely possibility that financial markets might have started to discount it in the equity price formation mechanism somewhat in advance.

Therefore, I finally settled on the beginning of January 1998, as this date maximises the likelihood of a ADCC-GARCH model of the bond index returns, it almost exactly splits the sample period in half and allows for the possibility that the correlations amongst euro-zone stock returns might have been affected by increased financial integration prior to the introduction of the new currency. Using the usual LR test statistic, reported at the bottom of Table 6.5, I therefore test Model 4 that allows for a structural break in 1998 against Model 2, the restricted model with no structural break. I can reject this restriction at the 0.020 significance level. Moreover, once I allow for the structural break, the restriction that the asymmetric component coefficient  $\theta$  is equal to zero (Model 5 against Model 4) cannot be rejected at the 5 percent level. The coefficient  $\theta$  is only marginally significant. Its size however is non negligible from an economic point of view. In particular, its point estimate is 45 times as large as the news reaction parameter  $\alpha$ .

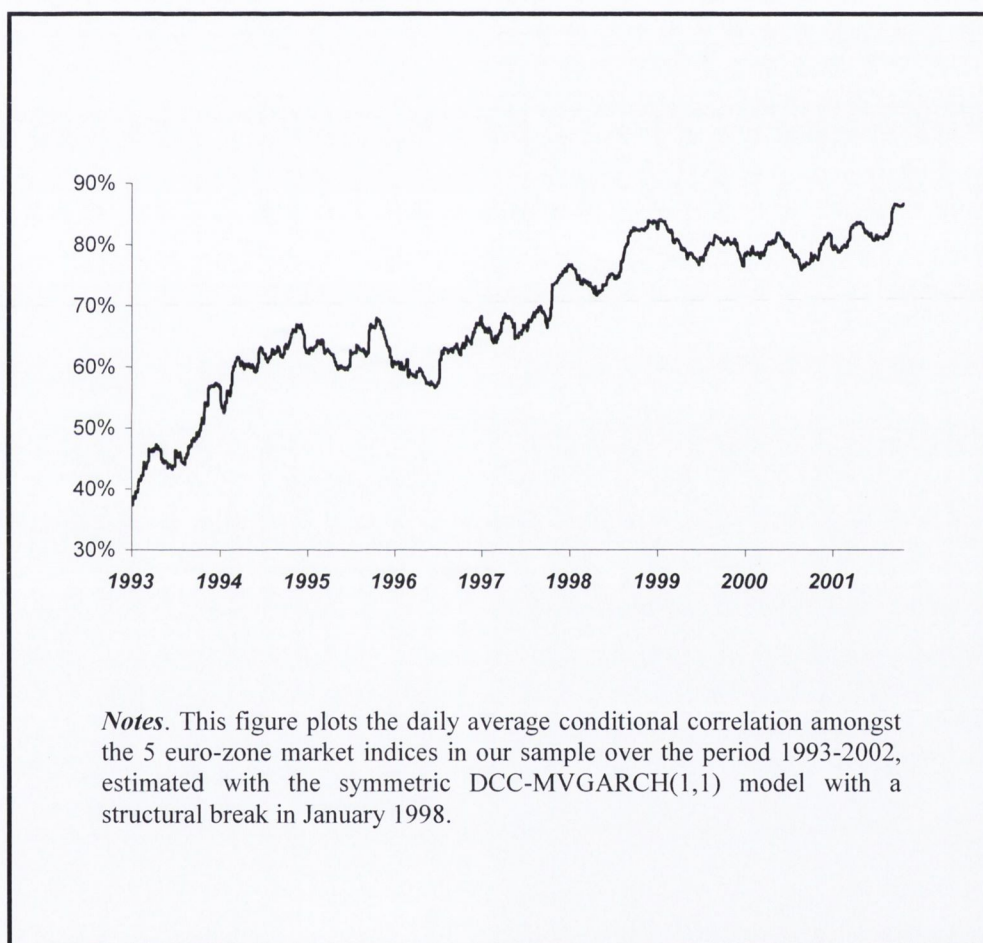


**Figure 6.2**  
**EMU Government Bond Yields**  
**Rolling Structural Break Dates Log-Likelihoods**



Since the estimate of the asymmetry parameter  $\theta$  is not very reliable due to the large sampling error, I conclude that the aggregate correlation between the 5 Euro-zone stock market indices and the Eurostoxx50 index is best modelled as a symmetric DCC-GARCH process with a structural break in its mean<sup>63</sup>. Figure 6.3 plots the market index average conditional correlation estimated with the symmetric Model 5, allowing for a structural break in January 1998.

**Figure 6.3**  
**DCC-MVGARCH Country Correlation**



<sup>63</sup> I also estimated each model with the *Eurostoxx50* index, and over the longer sample period 1992-2002, excluding the *MIB30* index (because its series starts a year later). I obtained very similar results in all cases, and these are not reported here for expositional clarity.



Turning to the correlation patterns at a more disaggregated level, the estimation results for selected specifications of the DCC-MVGARCH model with the 20 *stox* market sector indices and the 35 Datastream industry indices are reported in Table 6.6. The estimated asymmetric reaction coefficient  $\theta$  is relatively small in the case of the industry indices. For the market sector indices, instead, its point estimate is relatively large and its *t*-statistic is highly significant. However, the restriction that the symmetric model imposes on the asymmetric one cannot be rejected on the basis of a LR test (Model 1 vs. Model 2). In the case of the 42 individual stocks, as shown in Table 6.7, the estimated  $\theta$  is very small and the restriction that it is equal to zero (Model 1 against Model 2)<sup>64</sup> cannot be rejected at any conventional significance level. At the disaggregate industry and stock level, therefore, the type of asymmetric reaction to joint past good and bad news modelled using (6.7) is not the salient feature of the process followed by conditional correlations. At the sector level, however, it appears to be relatively more important.

The time series of the estimated symmetric average conditional industry, sector and stock return correlation is plotted in Figure 6.4. The plot for the asymmetric case, not reported, is very similar. Interestingly, visual inspection reveals that industry and market sector average correlations are much more stable (less volatile) than the 42 stocks average correlation. The stock portfolios with the most stable correlation structure are those that correspond to the Datastream classification into 35 industries, followed by the *estox* classification into 18 market sectors.

---

<sup>64</sup> I do not report estimates with a deterministic time trend because the estimation procedure did not converge.

**Table 6.6**  
**Conditional Sector and Industry Correlations**

<i>Panel A</i>					
<i>Model</i>	<i>Restriction</i>	<i>Coefficient</i>	<i>Coefficient estimate</i>	<i>T-Ratio</i>	<i>p-value</i>
<i>stoxx Market Sectors</i>					
1	$Q_1 = Q_2$	$\alpha$	.107	9.71	.000
	$\delta_{Trend} = 0$	$\beta$	.703	5.65	.000
		$\theta$	.122	5.92	.000
2	$Q_1 = Q_2$	$\alpha$	.091	17.14	.000
	$\delta_{Trend} = 0$	$\beta$	.988	456.49	.000
		$\theta = 0$	$\theta$	.000	
<i>Level 4 Datastream Industries</i>					
3	$Q_1 = Q_2$	$\alpha$	.007	9.29	.000
	$\delta_{Trend} = 0$	$\beta$	.919	55.04	.000
		$\theta$	.003	1.74	.080

<i>Panel B</i>						
<i>Unrestricted Model</i>	$\ln( \Sigma_{UR} )$	<i>Restricted Model</i>	$\ln( \Sigma_R )$	<i>LR Statistic</i>	<i>Significance Level</i>	<i>Restriction Rejection</i>
1	-14.9903	2	-14.9908	0.4562	0.499	No

$$LR = -T\ln(|\Sigma_{UR}|) - \ln(|\Sigma_R|) \sim \chi^2(q)$$

$$T = \text{number of observations (2,289)}$$

$$\Sigma_{UR} = \text{covariance matrix of the residuals of the unrestricted model}$$

$$\Sigma_R = \text{covariance matrix of the residuals of the restricted model}$$

$$\chi^2(q) = \text{Chi-Squared distributions with } q \text{ degrees of freedom}$$

$$q = \text{number of restrictions (} q = 1 \text{)}$$

*Notes.* Panel A of this Table reports the coefficients, *t*-statistics and *p*-values for the ADCC-MVGARCH model of conditional correlations amongst among 20 *stoxx* market sector indices and among 35 Level 4 industry Datastream indices for the period 1987-2004 estimated with weekly data. Variables and their coefficients are defined in the text. Panel B reports Likelihood Ratio (*LR*) test statistics and their significance level.



**Table 6.7**  
**ADCC-MVGARCH 42 Eurostoxx50 Stocks**

Panel A

<i>Model</i>	<i>Restriction</i>	<i>Coefficient</i>	<i>Coefficient estimate</i>	<i>T-Ratio</i>	<i>p-value</i>
1	$Q_1 = Q_2$	$\alpha$	.002	16.51	.000
	$\delta_{Trend} = 0$	$\beta$	.989	1222.29	.000
	$\theta = 0$				
2	$Q_1 = Q_2$	$\alpha$	.002	15.06	.000
	$\delta_{Trend} = 0$	$\beta$	.989	1214.20	.000
		$\theta$	.001	1.55	.121

Panel B

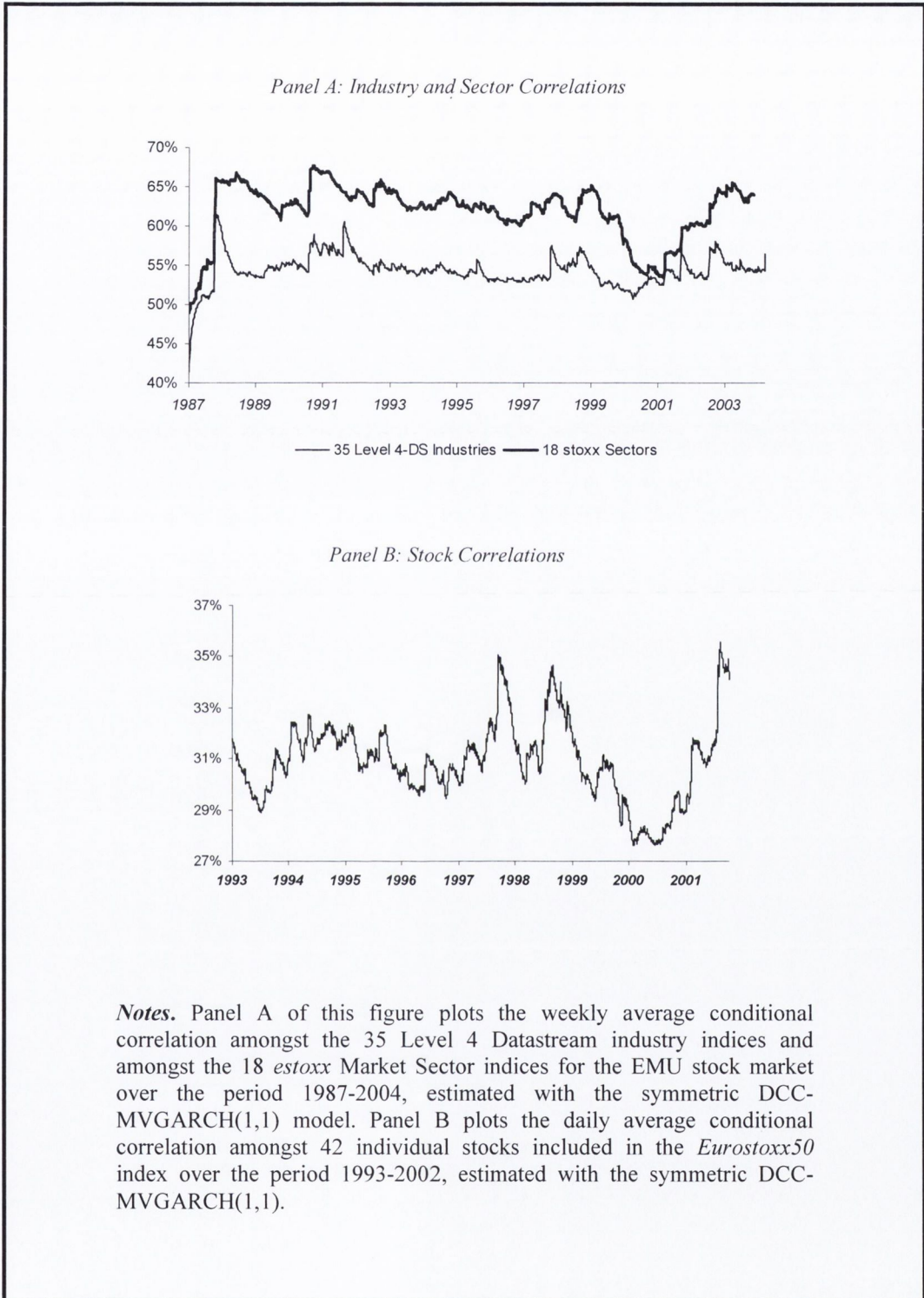
<i>Unrestricted Model</i>	$\ln( \Sigma_{UR} )$	<i>Restricted Model</i>	$\ln( \Sigma_R )$	<i>LR Statistic</i>	<i>Significance Level</i>	<i>Restriction Rejection</i>
2	-13.6466	1	-13.6474	1.7486	.186	No

$LR = T \ln(|\Sigma_{UR}|) - \ln(|\Sigma_R|) \sim \chi^2(q)$   
 $T = \text{number of observations (2,289)}$   
 $\Sigma_{UR} = \text{covariance matrix of the residuals of the unrestricted model}$   
 $\Sigma_R = \text{covariance matrix of the residuals of the restricted model}$   
 $\chi^2(q) = \text{Chi-Squared distributions with } q \text{ degrees of freedom}$   
 $q = \text{number of restrictions } (q = 1)$

*Notes.* Panel A of this Table reports the coefficients, *t*-statistics and *p*-values for the ADCC-MVGARCH model of conditional correlations amongst 42 stocks included in the Eurostoxx50 index over the sample period 1993- 2002. The data frequency is daily. Variables and their coefficients are defined in the text. Panel B reports Likelihood Ratio (*LR*) test statistics and their significance level.

**Figure 6.4**

**DCC-MVGARCH Industry, Market Sector and Stock Correlations**





As a specification check, I apply the Engle and Ng (1993) test in a multivariate setting. Originally, this test was designed as a diagnostic check for univariate volatility models and its aim is to examine whether there is residual predictability in squared standardised conditional errors using some variables observed in the past which are not included in the volatility model. Since multivariate variance-covariance models provide estimates of all the ingredients that are needed to compute the conditional portfolio volatility if asset weights are known, I can use my first and second step MV-ADCC and MV-DCC GARCH conditional volatility and correlation estimates to compute the conditional volatility and the conditional standardized residuals of an equally weighted portfolio. I can then apply the Engle and Ng (1993) test to the returns on the latter.

In particular, I apply the specification described in (3.33) that combines the sign bias test (that uses as regressors dummy variables  $I$  that take value 1 or 0 depending on whether the lagged residual is negative or positive) and the negative and positive size bias test (that use, respectively, lagged negative and positive standardised residuals as regressors,  $z_{t-1}^-$  and  $z_{t-1}^+$ ). As reported in Table 6.8, I can reject in every case the null of non-predictability of the squared standardised conditional residuals. As expected, the significance of the  $z_{t-1}^+$  term in (3.33) is unaffected by whether the  $\theta$  asymmetry coefficient in (6.7) is restricted to be equal to zero. The asymmetric specification of the DCC model of country level indices, with  $\theta \neq 0$  in (6.7), renders  $z_{t-1}^-$  insignificant. The latter however remains significant in the case of sector and industry indices.

In spite of conflicting evidence provided by the LR tests of the ADCC-MVGARCH against the DCC-MVGARCH, therefore, asymmetric influences of past innovations are nonetheless important. This suggests that, especially for sector and industry correlations, the ADCC-GARCH specification is unsuccessful at fully capturing the sources of the asymmetry in the data, which is probably of a non-linear nature and perhaps related to phenomena of contagion. I leave the difficult quest for a better specification for future research. The circumstance that the conditional correlation at

the market, industry and firm level appears to follow an asymmetric process of a different (yet unknown) type from the one modelled by the ADCC-GARCH specification is a noteworthy feature of my findings.

This result, as far as market indices are concerned, lies in partial contrast to those reported by Cappiello, Engle and Sheppard (2003). However, I am able to replicate their results with the same set of market indices, frequency and data period<sup>65</sup>. The difference between their results and mine, therefore, might be due to the different composition of the sample as correlations amongst EMU market indices appear to display a lower tendency to increase following joint past negative returns than those amongst markets outside the EMU. Another important issue is the different data frequency. Cappiello, Engle and Sheppard (2003) use only weekly data whereas I use both daily and weekly data and for the former the importance of the asymmetric correlation component is always lower.

---

<sup>65</sup> These results are not reported for brevity and because they exactly match results already published by Cappiello, Engle and Sheppard (2003) but they are available upon request.



**Table 6.8**  
**Diagnostic Tests**

Model	$\Gamma$ [sig.]	$z^-$ [sig.]	$z^+$ [sig.]	Chi-squared(3) [sig.]
Country Indices – Daily				
DCC	.068 [.394]	-.123 [.039]	-.142 [.058]	29.64 [.000]
ADCC	.099 [.396]	-.076 [.297]	-.169 [.048]	20.42 [.000]
Market Sector Indices – Weekly				
DCC	-.209 [.124]	-.567 [.000]	-.102 [.472]	37.60 [.000]
ADCC	-.217 [.110]	-.570 [.000]	-.107 [.452]	37.75 [.000]
Industry Indices Weekly				
ADCC	-.068 [.631]	-.325 [.004]	-.077 [.640]	11.46 [.009]

*Notes.* This Table reports the coefficients and  $p$ -values for a multivariate application of the Engle and Ng (1993) test. Variables and their coefficients are defined in the text.

#### **6.4. Summary and Conclusions**

The purpose of this Chapter is to contribute to the literature on the correlation dynamics in European equity markets. My main focus has been on the country level, but for comparison I also examined the behaviour of the correlations between the market sectors, between the industries and amongst a sample of large capitalization stocks. I applied both symmetric and asymmetric versions of the DCC-MVGARCH model of Engle (2001) and Engle and Sheppard (2002) and I extended them to allow for a deterministic time trend in the correlation process.

I confirm the dramatic correlation surge at the market index level reported by many previous contributions (such as the studies of Hardouvelis, Malliaropoulos and Priestley (2000), Fratzschler (2002) and Capiello, Engle and Sheppard (2003)). A structural break in the mean of average country-level correlation shortly before the introduction of the Euro accounts for both the strong persistence of its time series and its significant rise over the sample period. The structural break occurs somewhat in temporal proximity to the so called Asian crisis. I leave however for future research the investigation of the intriguing issues of how to disentangle the effects of the latter from the effect of the introduction of the Euro and of the extent to which the two effects might have interacted.

I do not find any evidence that the correlations among the sectors, the industries and the individual stocks have trended either up-wards or downwards over the sample period. In fact, they generally decrease around the period of the introduction of the Euro.

Thus, while the scope for country level diversification in EMU equity markets is become considerably more limited over time, the potential benefits from sector, industry and stock level diversification strategies are still substantial. In particular, they were unaffected by the process of monetary convergence culminated in the introduction of the Euro in 1998 and by the process of financial integration in the



EMU. The result regarding individual stock correlations is especially significant because the 42 stocks in my sample are heavily capitalized and they are traded in very liquid markets with low bid-ask spread. Brokerage fees to trade these stocks are also usually very low. Since their average correlation oscillates between 27 and 36 percent, these stocks offer the opportunity to achieve substantial diversification benefits incurring relatively low transaction costs.

Applying a multivariate generalization of the Engle and Ng (1993) specification test, I also find that the conditional correlation response to past positive and negative returns innovations in the EMU equity markets, while asymmetric, is not fully captured by the linear specification of the ADCC model. The tendency to rise following past negative returns is especially pronounced in sector indices and country level correlations. This asymmetric behaviour of correlations conditional on the sign of return realizations, coupled with a similar behaviour of volatilities, implies negative skewness of the multivariate distribution of the stocks and the indices under consideration and, in general, negative coskewness with a portfolio formed by the same stocks or indices. An investor that displays DIARA would seek to form portfolios that, for a given level of expected return and variance, display correlations that increase as little as possible following negative return realizations.

## **Chapter 7: The Time Series of Stock Market Returns**

### **7.1. Introduction**

In this Chapter, I discuss the equilibrium relation between first and second moments of the time series of stock market returns and I present a number of simple, yet novel analytical results on the relation between asset price determination, the equity premium and volatility. This Chapter is based on a paper forthcoming in the *Applied Financial Economics Letters*, Potì (2005a). In the next section, I compare and contrast the insights provided by the asset pricing and second moments literature reviewed in Chapter 2 and 3, respectively. In Section 7.3, using Campbell's (1991) unexpected return decomposition and the near unpredictability of dividend growth, I derive a simple relation that links the volatility of expected returns to conditional asset price volatility, and hence to the volatility of unexpected returns. In Section 7.4, I report the results of the application of this analytical framework to data on the S&P composite index for the period 1871-2003. Section 7.5 draws together the conclusions.

### **7.2. Discount Factor and Conditional Return Volatility**

According to the efficient market hypothesis (EMH) of Fama (1970), the difference between realised and expected asset returns depends only on changes to the available information set. For the EMH to hold, and if expected returns are constant as implied by the strict random walk model of asset prices, volatility should be caused solely by shocks to the available information set about expected asset cash flows. Alternatively, if we allow variation in expected returns, and for example we let the parameters of the SDF change over time, then realised volatility depends on shocks to the expected cash flow stream, on shocks to the expected rate of return, and on the covariance between the two types of shocks.



There are thus two sources of ex-post return volatility in an EMH framework. Changes to the expected cash-flow stream are the first volatility source. The second volatility source is given by changes in expected returns. Since dividends are not very volatile, the bulk of stock market volatility must be explained by volatility of the discount factor (the expected return). In other words, postulating high expected return volatility is one way to rationalise the perplexing empirical evidence that asset return volatility is much larger than the volatility of the dividend stream (one version of the so called ‘volatility puzzle’ formulated by Shiller (1981) and Campbell and Shiller (1988)). Asset pricing models provide a theoretical specification for volatility in discount rates.

A large body of literature has refined numerous models and techniques for estimating time-varying conditional asset price volatility, see for a review Bollerslev, Engle and Nelson (1994) or Chapter 3 of this thesis. Another large body of literature, initiated by Shiller’s (1981) seminal contribution on excess stock market volatility, has investigated at length the relation between asset price volatility and the volatility of expected fundamentals in order to check whether the former is explained by the latter as predicted by the rational valuation formula (RVF) under Muth’s (1961) rational expectation hypothesis (RE). Unfortunately, however, few studies, among these is the seminal paper of Campbell and Hentschel (1992), have made use of the sophisticated econometric tools developed to estimate conditional time varying volatilities to infer the implications of the RVF for discount rates volatility.

### 7.3. The Model

Consider Campbell’s (1991) unexpected return decomposition implied by the RVF with time varying expected dividends and returns:

$$r_t - E_{t-1}(r_t) = (E_t - E_{t-1}) \left( \sum_{j=0}^{\infty} \rho^j \Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^j r_{t+j} \right) \quad (7.1)$$

Here,  $r_t$  denotes continuously compounded rates of return,  $\rho = \frac{1}{1 + D/P}$ ,  $D$  and  $P$  are dividend and price at a an arbitrary point (e.g., taking the typical values over a sample of US equity data for the last 50 years,  $\rho = 0.96$ ),  $d_{t+j}$  is the log-dividend paid in  $t+j$  and  $r_{t+j}$  is the rate of return between  $t+j-1$  and  $t+j$ .

While dividends tend to grow over time, dividend growth appears to be nearly unpredictable. I therefore model (log) dividends as a random walk with a possibly non-zero deterministic drift and (log) dividend growth as a trend-stationary process with independently and identically distributed (i.i.d.) random residuals, i.e.  $d_{t+1} = d + d_t + e_{t+1}$  and  $\Delta d_{t+1} = d + e_{t+1}$ . Empirical evidence shows instead that returns are to some extent predictable and heteroskedastic. Therefore, I model returns as a possibly mean reverting<sup>66</sup> persistent process, i.e.  $r_{t+1} = \alpha + \beta r_t + \varepsilon_{t+1}$ . Here,  $\varepsilon_{t+1}$  are serially uncorrelated but possibly heteroskedastic residuals. See, for a review of the empirical evidence and a discussion, Cochrane (2001) or Section 1.3. On the basis of these assumptions about the data generating process of dividend growth and returns, (7.1) can be rewritten as follows<sup>67</sup>:

<sup>66</sup> Sometimes, the term mean-reversion is used to refer to the behaviour of variables with negative autocorrelation whereas persistence is used where autocorrelation is positive. My use of the term "mean reversion" encompasses both.

<sup>67</sup> From (7.1), the first term  $e_t$  of (7.2), is the difference between the expectations of the 'discounted' dividend stream conditional on information available at  $t$  and  $t-1$ . Clearly, the two streams differ only by the amount represented by the dividend innovation in  $t$ , i.e.

$$(E_t - E_{t-1}) \left( \sum_{j=0}^{\infty} \rho^j \Delta d_{t+j} \right) = \rho^0 e_t = e_t$$

Turning to the second term in the first line of (7.2), it can be rewritten as follows:

$$(E_t - E_{t-1}) \left( \sum_{j=1}^{\infty} \rho^j r_{t+j} \right) = (E_t - E_{t-1}) \left( \sum_{j=1}^{\infty} \rho^j \beta^{j-1} r_{t+1} \right)$$

Dividing and multiplying the right-hand side by  $\rho$ :

$$(E_t - E_{t-1}) \left( \sum_{j=1}^{\infty} \rho^j \beta^{j-1} r_{t+1} \right) = (E_t - E_{t-1}) \rho \left( \sum_{j=1}^{\infty} \rho^{j-1} \beta^{j-1} r_{t+1} \right)$$

Finally, recognising that the term in bracket is a power series in  $\rho\beta$  times the return in  $t+1$  gives the expression in the second line:

$$(E_t - E_{t-1}) \left( \sum_{j=1}^{\infty} \rho^j r_{t+j} \right) = (E_t - E_{t-1}) \rho \left( \frac{1}{1 - \rho\beta} r_{t+1} \right) = (E_t - E_{t-1}) \left( \frac{\rho}{1 - \rho\beta} r_{t+1} \right)$$



$$\begin{aligned}
r_t - E_{t-1}(r_t) &= e_t - (E_t - E_{t-1}) \left( \sum_{j=1}^{\infty} \rho^j r_{t+j} \right) \\
&= e_t - \frac{\rho}{(1 - \rho\beta)} (E_t - E_{t-1}) r_{t+1}
\end{aligned}
\tag{7.2}$$

Taking the unconditional expectation of the square of both sides of (7.2), the conditional return volatility of conditionally unexpected returns can be approximated as follows:

$$E[r_t - E_{t-1}(r_t)]^2 \cong \frac{\rho^2}{(1 - \rho\beta)^2} E[(E_t - E_{t-1}) r_{t+1}]^2
\tag{7.3}$$

This approximation neglects both the second moment of  $e_t$  and its cross-moment with  $(E_t - E_{t-1}) r_{t+1}$  because dividends are not very volatile and do not forecast returns. Solving (7.3) for the unconditional second moment of discount rates and using the shorthand notation  $\sigma^2$  for the unconditional variance of conditional residuals:

$$\begin{aligned}
E[(E_t - E_{t-1}) r_{t+1}]^2 &\cong E[r_t - E_{t-1}(r_t)]^2 \frac{(1 - \rho\beta)^2}{\rho^2} \\
&\cong \sigma^2 \frac{(1 - \rho\beta)^2}{\rho^2}
\end{aligned}
\tag{7.4}$$

Equation (7.4) relates the amount of discount factor volatility to unconditional volatility of conditional unexpected returns, under the assumption that dividend growth is not very volatile and does not forecast returns.

The assumption that dividend volatility is negligible is a limiting case that it is useful to consider in order to focus on the relation between the volatility of returns and discount factors. This somewhat extreme assumption, however, is unpalatable from an empirical point of view. Thus, squaring both sides of (7.2), taking their

unconditional expectations, allowing for dividend volatility but still assuming that dividends do not forecast returns, we can write:

$$E[r_t - E_{t-1}(r_t)]^2 \cong E(e_t^2) + \frac{\rho^2}{(1 - \rho\beta)^2} E[(E_t - E_{t-1})r_{t+1}]^2 \quad (7.5)$$

Solving (7.5) for the discount factor volatility yields another relation between unconditional discount factor volatility and the unconditional volatility of conditional unexpected returns, under the assumption that dividend growth does not forecast returns:

$$\begin{aligned} E[(E_t - E_{t-1})r_{t+1}]^2 &\cong \left\{ E[r_t - E_{t-1}(r_t)]^2 - E(e_t^2) \right\} \frac{(1 - \rho\beta)^2}{\rho^2} \\ &\cong (\sigma^2 - \sigma_d^2) \frac{(1 - \rho\beta)^2}{\rho^2} \end{aligned} \quad (7.6)$$

Here,  $\sigma_d = \sqrt{E(e_t^2)}$  denotes dividend growth volatility. The upshot of (7.4) and (7.6) is to clarify how much variability we need in discount rates (or expected rates of return) under alternative assumptions about the magnitude of dividend growth volatility, and the unconditional volatility of unexpected returns, given returns predictability, unpredictable dividend growth, and knowledge of the typical dividend-price ratio. This result provides a useful criterion by which competing asset pricing models can be evaluated. If the asset pricing model under consideration does not produce variability of discount rates of the required magnitude, it can be discarded on grounds that it will not manage to explain observed conditional stock volatility. This criterion can also be used to evaluate competing reduced form econometric specifications. If a particular specification does not produce the required variability of discount rates, given estimated conditional return volatility or the estimated unconditional volatility of unexpected returns, it cannot be considered the reduced form of an admissible (under the RVF) asset pricing model. Cochrane (2001, 2005), provides an alternative yet fully equivalent representation of (7.6),



derived from a VAR model of a latent expected return variable, returns, and unpredictable dividend growth. The latent expected return variable is assumed persistent for identification purposes. In Cochrane's (2001, 2005) representation, the volatility of expected returns, i.e. the right-hand side of (7.6), is given by the dividend-price ratio volatility times a function of the amount of predictability in the latent expected return variable. Relative to the representation derived by Cochrane (2001, 2005), (7.6) is more straightforward in that it relates the volatility of expected returns to familiar quantities such as return and dividend volatility and the persistence of returns.

The return persistence parameter  $\beta$  will need to be estimated or calibrated to fit the data. If  $\rho\beta = 1$ , discount rate volatility must be zero. However, since typically  $\rho < 1$ , this would imply that  $\beta > 1$  and, therefore, that returns are explosive. With  $\beta = 0$  and neglecting dividend volatility, the required unconditional variability of discount rates is larger than the unconditional volatility of returns. In general, neglecting dividend volatility, the required variability of discount rates is lower than the volatility of returns when  $0 < \beta \leq 1$ , i.e. some persistence in returns and thus predictability help explain unexpected return volatility. In particular, when returns are so persistent that they follow a random walk, i.e.  $\beta = 1$ , very little discount rates variability is enough to account for the volatility of returns.

#### **7.4. Empirical Application**

The unconditional variance  $\sigma^2$  of conditional residuals in (7.4) can be estimated by a traditional ARCH or GARCH model, see for example Engle (1982) and Bollerslev (1986), as it is simply the value to which the conditional variance of returns converges in the long run. A GARCH-M model, such as the specification proposed by Engle, Lilien and Robins (1987), would also provide estimates of the time varying conditional trade-off between expected return and risk. To be more specific, consider the following GARCH-M model of returns:

$$E_{t-1}(r_t(\theta)) = g[\sigma_{t-1}^2(\theta), \theta] \quad (7.7)$$

Here, a non-zero derivative of  $g$  with respect to the first element of its argument allows the conditional mean to be an explicit function of conditional variance  $\sigma_{t-1}^2 = E_{t-1}(\varepsilon_t^2(\theta))$ ,  $\varepsilon_t(\theta) = r_t - E_{t-1}(r_t(\theta))$  is a discrete time stochastic error process with conditional mean and variance parametrized by the finite dimensional vector  $\theta \in \Theta \subseteq R^m$ ,  $\theta_0$  denotes the true value of the parameter vector. In (7.7), the conditional mean of the residuals is  $E_{t-1}(\varepsilon_t(\theta_0)) = 0$  but their conditional variance  $E_{t-1}(\varepsilon_t^2(\theta_0)) = \sigma_{t-1}^2$  is allowed to be time varying. The unconditional variance of the conditional residuals is  $E(\varepsilon_t^2(\theta_0)) = \sigma^2$ .

A possible empirical specification of (7.7) is represented by a GARCH(1,1)-M model with a mean equation that includes, in addition to the conditional variance term, a constant and one lag of the stock market return:

$$\begin{aligned} E_{t-1}(r_t) &= \alpha + \alpha_1 \sigma_{t-1}^2 + \beta r_{t-1} \\ \sigma_t^2 &= \alpha_2 + \alpha_3 \varepsilon_{t-1}^2 + \alpha_4 \sigma_{t-1}^2 \end{aligned} \quad (7.8)$$

I estimate conditional annualized volatilities of the S&P Composite Index<sup>68</sup> over the period 1871-2003 using a GARCH(1,1) specification with a first-order autoregressive model for the conditional mean. I therefore estimate (7.8) with the restriction that  $\alpha_1 = 0$ . The ARCH specification allows for more efficient estimates of the parameters of (7.8), including the crucial persistence parameter  $\beta$ , but the usual trade-off between efficiency and robustness applies. The point estimate of  $\beta$  is about 0.35 and the unconditional volatility estimate, given by  $\sigma = \frac{\alpha_2}{1 - \alpha_3 - \alpha_4}$ , is about 14 percent.

---

<sup>68</sup> I thank Professor Robert Shiller for making this data available on his website, and for helpful discussion.



With  $\beta = 0.35$  and this estimate of  $\sigma$ , the annualized volatility of monthly discount rates implied by (7.4) is just above 10 percent. Thus, neglecting dividend growth volatility, we need volatility in expected returns as large as 10 percent to explain the conditional return volatility over the period 1871-2003. The conditional stock market volatilities estimates from (7.8) and the implied discount factor variability bound are plotted in Figure 7.1.

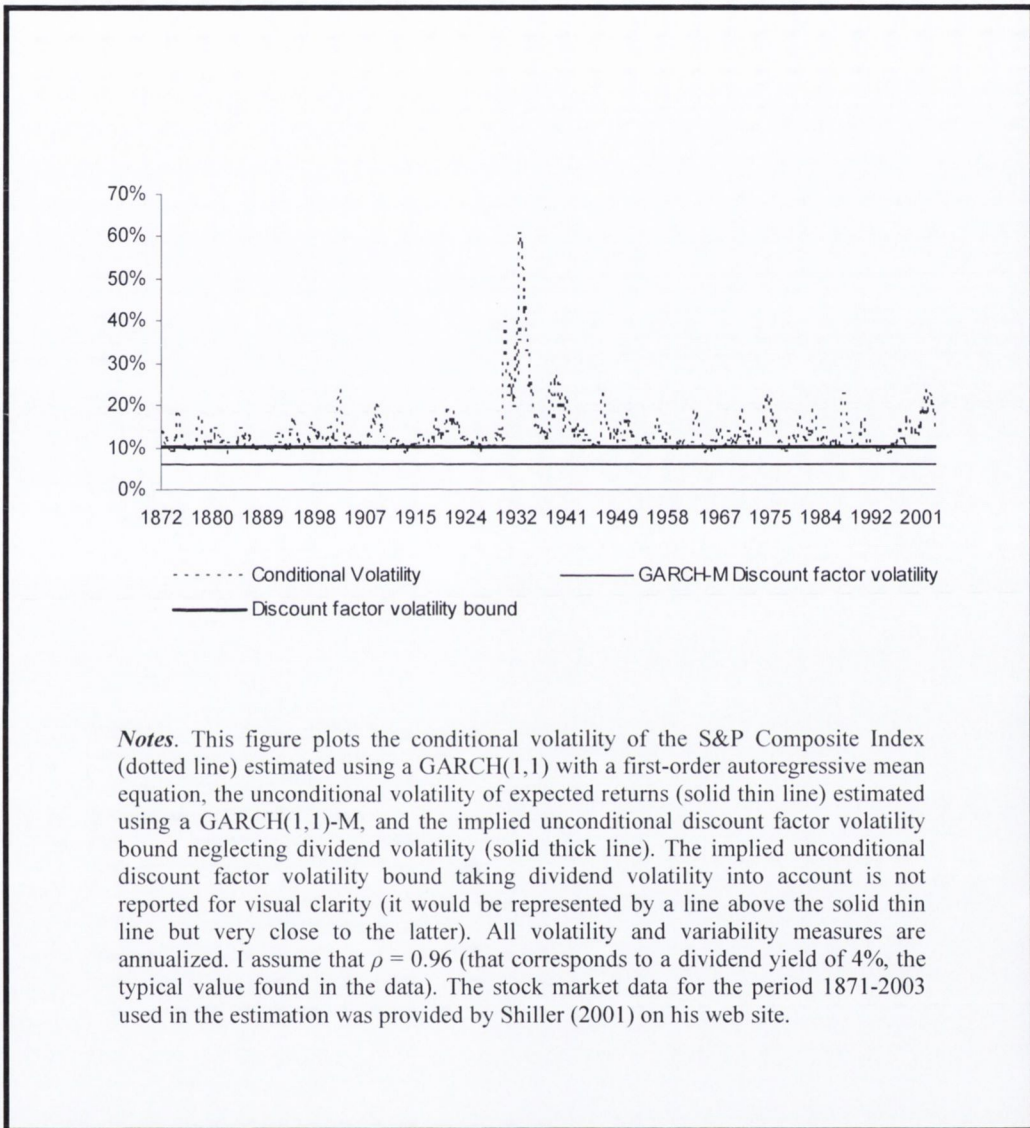
Dividend growth volatility in the 1871-2003 period is about 4 percent per annum. Since dividend growth is assumed to be i.i.d., its conditional and unconditional volatility are the same. Using (7.6), the discount rate volatility bound is about 7.1 percent with  $\beta = 0.35$ ,  $\sigma^2 = 14$  percent and 4 percent dividend growth volatility. Thus, taking this dividend growth volatility into account, we need a lower discount factor volatility to explain conditional return volatility.

The calculated values for the discount rate variability bound can be used to evaluate the empirical performance of competing asset pricing models. Sharpe (1964) and Lintner's (1965) static CAPM is very easily discarded, as it implies an obvious violation of the variability bound from below, even taking dividend growth volatility into account. A GARCH-M, broadly along the lines of Merton's (1973) intertemporal CAPM (ICAPM) but with no state variable to proxy for changes in the future opportunity set, produces a yearly volatility of the discount rate of 6.15 percent, as shown in Figure 7.1. This is computed as the sample annualized volatility of conditional mean returns estimated using (7.8) with  $\beta = 0$  but without the restriction that  $\alpha_1 = 0$ . The estimated variability of discount rates is still too low to satisfy the bound. It is much lower than the 10 percent bound under the assumption that dividend growth is negligible and just below the 7.1 percent bound that corresponds to 4 percent dividend growth volatility. Other models that price assets on the basis of their covariance with little volatile fundamental economic variables, such as aggregate consumption as in Lucas' (1978) Consumption CAPM, will struggle even more to reach the bound from below, unless we allow for additional sources of conditional discount factor volatility.

The estimated lower bound on the volatility of expected returns has interesting implications for the quasi-concavity of the expected utility implied by the SDF estimates reported in Table 4.2 (in Chapter 4). The unrestricted estimates imply a convex SDF, convex marginal utility, and an S-shaped utility function. The flexus of marginal utility occurs at a level of quarterly excess return of around 2.5 percent, or 10 percent annualized. Assuming a 6 percent risk free rate, this corresponds to a 16 percent annual return. Thus, returns above this threshold are in the risk seeking region. To assess how likely these returns are, it is useful to construct a confidence interval around the mean return, about 12 percent in sample (about 8 percent real excess return plus 6 percent nominal risk free rate). Given my estimated 7.1 percent expected return volatility lower bound, the sampling error of the mean, with a sample of over 100 annual observations, is at least about  $0.71 = 7.1/10$ . This means that a 2-tailed 2-standard deviation interval for the mean return is  $\{12 - 1.42 \text{ percent}, 12 + 1.42 \text{ percent}\}$ , i.e. about  $\{10.8, 13.2 \text{ percent}\}$ . As return annual volatility is about 14 percent, the 16 percent annual return threshold of the risk seeking region is within 1 standard deviation from either the lower or upper end of the confidence interval for the mean. This back-of-the-envelope calculation suggests that returns in the risk seeking region cannot be considered low probability events. The implication for the analysis in Chapter 4 is that it is unlikely that expected utility be quasi-concave.



**Figure 7.1:  
Conditional Return Volatility vs. Discount Factor Variability**



*Notes.* This figure plots the conditional volatility of the S&P Composite Index (dotted line) estimated using a GARCH(1,1) with a first-order autoregressive mean equation, the unconditional volatility of expected returns (solid thin line) estimated using a GARCH(1,1)-M, and the implied unconditional discount factor volatility bound neglecting dividend volatility (solid thick line). The implied unconditional discount factor volatility bound taking dividend volatility into account is not reported for visual clarity (it would be represented by a line above the solid thin line but very close to the latter). All volatility and variability measures are annualized. I assume that  $\rho = 0.96$  (that corresponds to a dividend yield of 4%, the typical value found in the data). The stock market data for the period 1871-2003 used in the estimation was provided by Shiller (2001) on his web site.

## 7.5. Summary and Conclusion

In this Chapter, I have derived the relation between the unconditional volatility of conditionally unexpected and expected returns, given conditional return volatility, near unpredictability of stock dividends, the typical dividend-price ratio, and return persistence. This result can be used to place a bound on the volatility of conditional expected returns that, under the RVF, any econometric model must produce to be the empirical specification of an admissible asset pricing model. I discussed how the expected return conditional volatility bound can be used to evaluate the empirical admissibility of Sharpe (1964) and Lintner's (1965) static Capital Asset Pricing Model (CAPM), of a GARCH-M based on Merton's (1973) CAPM and of Lucas' (1978) CCAPM. A useful extension of this work would be the evaluation of a larger set of competing theoretical and econometric specifications.

One limitation of the derivation is that returns can only mean revert to a constant long run level. This is implied by the proposed autoregressive specification of the return process. A useful extension would be the multivariate generalization of (7.6) and, in particular, the derivation of the relation between conditional return volatility, conditional dividend volatility and unconditional expected return volatility when the latter are driven by a persistent, slow moving state variable. This would amount to a relatively simple generalization of the vector autoregressive model (VAR) presented by Cochrane (2001) to explain return predictability. To render this model able to generate conditional return volatility, it suffices to allow for heteroskedasticity in the process followed by the errors of the dividend growth and/or of the state variable. In particular, their second moments should be made depend, in a non trivial fashion, on the past of the processes themselves. Campbell and Hentschel (1992) already did this for the dividend process. I leave these developments for further research.



## **Chapter 8: Implications, Limitations and Future Work**

### **8.1. Introduction**

In this Chapter I summarize my main findings and I discuss their implications for asset pricing and investment management. I then point to the limitations of the thesis. Finally, I point to directions for future research. I start, in the next Section, by reviewing the main findings reported in this thesis and the important analytical results. Section 8.3 discusses their implications. In Section 8.4, I discuss certain epistemological issues that arise in connection with a number of tests discussed in this thesis. In Section 8.5, I then outline the main limitation of this study, I suggest possible extensions and I highlight opportunities for future research. The final Section presents some final remarks and draws together the main conclusions.

### **8.2. The Main Findings Restated**

My thesis contributes to depict a representation of the multivariate distribution of stock returns where the relations between moments and their dynamics are important in explaining both the time series behaviour of stock returns and their cross-sectional differences. Chapter 4 expanded the extant evidence that coskewness helps explain the cross section of US returns. In particular, the unconditional pricing implications of a conditional stochastic discount factor model quadratic in the stock market excess return explain about 30 percent of the cross-section of the average returns on 30 US industry indices. I interpreted this finding as evidence that the correlation with systematic second moments realizations (the squared market returns) helps explain the differences in first moments (expected returns) across stocks. However, I also confirmed that, as reported by Dittmar (2002) and Post, Levy and van Vliet (2003), the estimated stochastic discount factor violates the concavity requirement. This renders problematic its interpretation as the stochastic discount factor implied by the 3M-CAPM. I showed that a conditional specification that uses Lettau and Ludvigson's (2001) proxy for the consumption-wealth ratio suffers from this

problem too. Chapter 5 added to the extant evidence that there is a positive relation between aggregate returns and systematic risk for the broad EMU equity market. I confirmed that market and average idiosyncratic variance, as already documented by Goyal and Santa Clara (2003) and by Guo (2003) using US data, predict market-wide returns. I demonstrated, however, that the sign of the relation between market variance, idiosyncratic variance and market returns implies that average correlation predicts market returns and thus that there is a positive relation between a component of systematic risk and aggregate returns. Chapter 7 clarified that it takes a substantial volatility of discount factors (conditional first moments) to explain second moments of unexpected returns. In particular, the discount factor volatility must be about 7 percent per annum to explain the volatility of conditionally unexpected returns on the US stock market over the period 1871-2003.

My thesis also highlighted the many relations between second moments and the extent to which the stylized empirical features of the dynamic behaviour of systematic and idiosyncratic volatility and correlations in US equity markets are to be found also in EMU equity markets. To summarize the behaviour of the correlations between a large number of stocks, I introduced a synthetic average correlation measure. This corresponds to the level of correlation that, if assumed to hold for all the assets in a portfolio, would yield the same portfolio volatility as the full correlation matrix. To facilitate the construction of the average correlation time series, I showed that the latter can be approximated as the ratio of market to squared average total volatility. I then studied the average correlation time series to infer the salient features of the common dynamics of stock correlations.

As reported in Chapter 6, there has been a surge in the correlations between EMU markets. For the main EMU stock markets, this is often close to 100 percent. This confirms the findings reported by many previous contributions (such as Hardouvelis, Malliaropulos and Priestley (2000), Fratzschler (2002), Capiello, Engle and Sheppard (2003)). Using the DCC-MVGARCH model of Engle (2001) and Engle and Sheppard (2002), I find that this dramatic surge, in spite of the significant degree



of persistence of conditional correlations, does not imply a non stationary long-run behaviour but it is explained instead by a once off structural break shortly before the introduction of the Euro. This confirms the results reported by Cappiello, Engle and Sheppard (2003) and is consistent with the rise in volatility spillovers noticed by Baele (2002). In contrast to this, however, there appears to be no structural break in the conditional correlation process at the firm and industry level, as shown in Chapter 5 and 6. The value-weighted average stock correlation is a fast moving series that tends to mean revert quickly to a 20 percent long-run mean after a shock. This relatively low mean of average stock correlation implies a correspondingly low explanatory power, or  $R^2$ , of the market model. The  $R^2$  of the market model for the average stock implied by this level of average correlation is about 4 percent, thus roughly half the value reported by CLMX (2001) for US stocks. I find strong evidence of asymmetry in the conditional correlation response to past return innovations of different sign and in general of asymmetry in the multivariate distribution of asset returns. I also find, however, that the type of asymmetry captured by the ADCC-GARCH models is not the most important source of asymmetry in the distribution of asset returns, likely of a non-linear nature.

The variance of both the average European stock and of the EMU market portfolio has substantially increased over time and a large portion of this increase is explained by a long-run trend. Market volatility and the idiosyncratic volatility of the typical stock tend to increase by about 7.5 and 10 percent, respectively, every 10 years. European stocks, therefore, have indeed become more volatile. Idiosyncratic volatility accounts for the main portion of the variance of the typical stock. EMU variance series are best forecast by market variance, whereas US variance series are best forecast by average idiosyncratic variance.

### **8.3. Implications**

There are two main sets of implications of my findings. The *first* relates to asset pricing. The *second* relates to the consequences for portfolio management of the rise

in average idiosyncratic volatility, of the dynamics of average correlation and of the coskewness. I will now discuss these implications in turn.

### *Asset Pricing Implications*

The beta-pricing representation in (4.1) of the 3M-CAPM, estimated using the 30 Fama and French US Industry portfolios and the CRSP index as a proxy for the market portfolio, implies a ‘coskewness puzzle’. The puzzle arises because, while the  $\lambda_{1,t}$  and  $\lambda_{2,t}$  parameter estimates fit the cross section of industry returns relatively well, they imply risk seeking over gains and thus a non-concave utility function. A similar puzzle arises from the estimates of the beta-gamma representation in (2.54), even allowing for conditional time-variation in the shape of the utility function. Given the parameters of the SDF implied by these estimates, the market portfolio is not necessarily efficient for the representative investor. In turn, if the market portfolio is inefficient, the 3M-CAPM does not hold.

Pending the investigation into the theoretical explanation of these findings, the interesting question is then whether we should price assets based on expected returns that reflect a coskewness premium and, in particular, the large coskewness premium  $\lambda_{2,t}$  implied by the unrestricted quadratic SDF specification. It is clear that, as long as we do not have an equilibrium asset pricing model that can account for this large coskewness premium, we cannot strictly consider coskewness a risk measure. However, since  $\beta_{i1,t}$  and  $\beta_{i2,t}$  are the factor loadings of a multifactor model that explains the cross-section of industry returns relatively well, we might draw pricing implications for other assets based on no-arbitrage arguments.

Thus, in the spirit of statistical multifactor models, we might price assets that are spanned by the industry portfolios using the unrestricted estimates of  $\lambda_t$ . In particular, in order to avoid extending the asset pricing implications of the unrestricted estimate of (4.1) to variation of stocks returns not spanned by the market excess return and its square, we should price only stocks that are spanned by those



industries that are priced accurately, such as Games, Constructions, Autos, Carry, Mines, Telecom, Paper and Wholesale. We might also add more factors to increase the portion of the stock variation explained by the multifactor model. We might further assume, as in Ross' (1976) APT, that the residual variation is of an idiosyncratic nature and thus use the multifactor model to price diversified portfolios of stocks. Yet, to apply this multifactor model to non diversified portfolios and thus to the constituent assets we would have to rule out the possibility that the price of the idiosyncratic residuals is high. One way to achieve this result is to rule out 'good deals' by bounding the volatility of the SDF, as suggested by Cochrane and Saà-Requejo (2000) and Cochrane (2001)<sup>69</sup>.

### *Investment Strategy Implications*

A conventional rule of thumb, based on Bloomfield, Leftwich and Long (1977), suggests that a randomly chosen portfolio of 20 stocks produces most of the reduction in idiosyncratic risk that can be achieved through diversification. However, as discussed by CLMX (2001), the higher the average idiosyncratic variance, the larger the number of stocks needed to achieve a relatively complete diversification, given a random portfolio selection strategy. In Panel A of Figure 8.1, I report the residual portfolio idiosyncratic volatility as a function of the number of stocks included in equally-weighted portfolios formed by drawing randomly from my stock sample for various levels of average idiosyncratic risk at different points in time. To reduce idiosyncratic volatility to 5.0 percent it took 166 stocks in 2003, 43 stocks in 1989 but just 35 stocks in 1974. It is worth noticing that most of the increase has taken place in the second half of the sample period. CLMX (2001)'s findings are similar. They report that a residual portfolio idiosyncratic volatility as low as 5 percent required 50 US stocks in the period 1986-1997 whereas it would

---

<sup>69</sup> I explored this possibility in a recent working paper of mine, Poti (2005c), that is available for download from the Social Science Research Network website (details on its web address are in the Bibliography). Please note, however, that the section that contains the implementation of this approach is still preliminary and incomplete.

have taken only roughly 20 stocks in the period 1974-1985<sup>70</sup>. Panel B of Figure 8.1 illustrates the combined effect of changes in average idiosyncratic variance and average correlation on the extent to which random diversification strategies can reduce the fraction of an equally-weighted portfolio variance represented by diversifiable idiosyncratic variance.

The higher the correlation among stock returns, the lower the potential benefit from diversification. Thus, overall, my results on the correlation dynamics in EMU markets suggest that fund managers should think through the full ramifications of seeking cost-effective diversification in the Euro-zone area by adopting the passive strategy of investing in market indices rather than a selection of stocks from the whole supra-national market. In particular, because of the rise in correlations among national stock markets indices, the stochastic components of the latter are now expected to behave almost identically (with conditional correlations being close to 100 percent). This suggests that there is little expected benefit from strategies that diversify across Euro-zone market indices, although diversification across stocks remains useful. In particular, the very low level of average correlation in the equally-weighted case suggests that diversification can be an important source of improvement in the portfolio risk-return ratio (even though the full benefit in terms of variance reduction will not be available to the average investor since average correlation in the value-weighted case is substantially higher).

While awareness of the degree of predictability of asset returns leads to a better assessment of the level of risk at different investment horizons, an understanding of the role of the non-diversifiable component of the moments of odd order, primarily asset co-skewness and portfolio skewness, makes it possible to better address the investment decision from a strategic point of view. From this perspective, the investor is faced not only with the mean-variance trade-off, but also with the choice between portfolios that tend to perform better or worse as aggregate correlations and

---

<sup>70</sup> It would also have taken about the same number of stocks in the earlier 1962-1973 period.

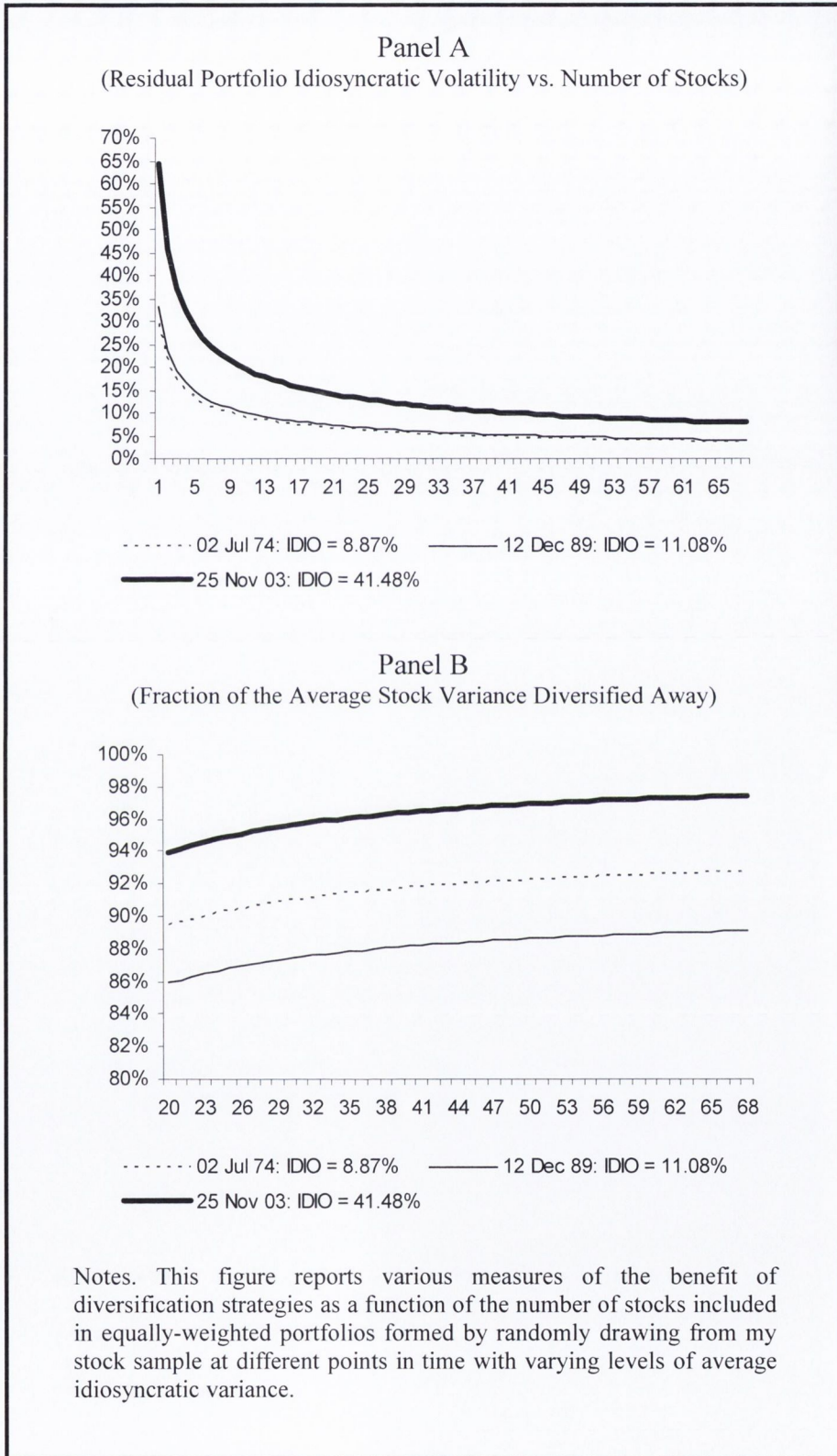


average volatilities increase and, therefore, as the overall market portfolio becomes riskier.

#### **8.4. Epistemological Note**

Any empirical specification of theoretical asset pricing models that derives the conditioning information set from the time-series of ex post excess-returns relies on the assumption that the latter contain sufficient information to infer the estimated risk-return relation. This problem, *mutatis mutandis*, is similar to the one highlighted by Roll (1977) with regard to tests of the CAPM. In particular, Roll's critique points out that any test of the CAPM is a joint test of the theoretical model and of the efficiency of the market portfolio. In an inter-temporal framework, whenever the conditional risk-return relation is estimated using ex-post returns as the conditioning information set, this leads to joint tests of the theoretical asset pricing model and the auxiliary assumption that the information set observable by the econometrician is sufficient to infer the true return data generating process. Hansen and Richard (1987) argue a similar point. The implication of this line of reasoning is that empirical tests of the risk-return relationship must be taken with caution and inferences should be frequently checked against new evidence that expand the information set available to the econometrician (e.g. to check, for example, that the model parameters estimates are not biased by a small sample problem related to high impact-low probability events, the so called peso problem).

**Figure 8.1**  
**Random Diversification and Idiosyncratic Volatility**





## 8.5. Limitations of the Analysis and Future Work

Predictability and time varying risk premia likely reflects a premium for holding macroeconomic risk associated with the business cycle and for holding assets that do poorly in times of high volatility and financial distress. Therefore, they seem to be closely related to the issue of asymmetry and thick tails in the multivariate return distribution and hence to asset coskewness. The exploration of the link between aggregate idiosyncratic risk, higher moments and asymmetries of the multivariate distribution of asset returns opens fascinating yet challenging possibilities for future research. For example, further research might suitably expand the set of conditioning variables to better model variation in the utility function parameters and might use a more meaningful proxy for the market portfolio of all risky assets. This, beside improving the fit of the model, might lead to a 3M-(C)CAPM specification with parameter estimates that do not violate RA and NIARA. Moreover, since I find that average idiosyncratic volatility is persistent and predicts aggregate market returns, it could be used as a scaling variable to proxy for background risk. Including average idiosyncratic variance amongst the set of conditioning variables used in a (C)CAPM empirical specification could be particularly useful because it is more readily available and it can be estimated with less delay than Lettau and Ludvigson's (2001) consumption-wealth ratio. This would bring together my research on aggregate idiosyncratic risk and on conditional asset pricing with higher moments. I leave these developments, however, for future research.

Turning to the 'coskewness puzzle', its solution requires a theory that predicts a stochastic discount factor quadratic in the market return without implying that the market portfolio is efficient. We might appeal to Harrison and Krepp's (1979) theorem to motivate the stochastic discount factor representation of the asset pricing problem without requiring that the market portfolio maximizes investors' expected utility. Recall that this theorem states that, given free portfolio formation and under the law of one price, there exists an  $m_{t+1}$  such that, for every payoff  $x_{t+1}$ ,  $p_t = E_t(m_{t+1}x_{t+1})$ . This approach, however, rises the problem of motivating why (2.60)

specifies  $m_t$  as a function of the market excess return and its square. Alternatively, we might specify individual utility functions that exhibit DIARA and then determine equilibrium prices without imposing restrictive assumptions such as investors' homogeneity and market completeness or the equivalent representative investor assumption. The interesting question then becomes why the market return and its square should be good proxies for aggregate marginal utility growth even though the market portfolio is not necessarily efficient for the representative investor. I leave the investigation of these issues for further research.

The exploration of asset pricing frameworks that explicitly take uninsurable idiosyncratic risk, non-symmetric return distributions and time varying risk premia into account opens fascinating possibilities for the formulation of richer portfolio and investment advice than the standard, traditional recommendation to hold the market portfolio implied by the static CAPM under the complete capital market assumption. For example, as summarized by Cochrane (1999), there are strategies that result in high average returns without large betas, i.e. with no strong tendency for the strategy's returns to move up and down with the market as a whole. The returns on these strategies cannot be rationalized on the basis of the traditional CAPM. The investigation of whether these abnormal returns can be explained by models that include the squared market return and aggregate idiosyncratic volatility as additional factors would be a useful extension of this thesis.

Turning to the dynamic behaviour of second moments, my study of aggregate, industry and firm level volatility and correlation in EMU markets could be extended in several directions. For example, while the information carried by low frequency series is important in many asset allocation problems, high frequency series are also important in trading and risk management applications more focussed on the short run behaviour of stock returns. Thus, a useful extension of the study of EMU volatilities and correlations would be to construct higher frequency series from daily or even intraday returns. Care should be taken then to overcome the problems associated with asynchronous trading in the various European markets. Also, it



would be useful to construct the average industry-level variance series directly from stock returns, instead of using Datastream industry indices. Further investigating the role of idiosyncratic volatility in asset pricing, with special regards to whether it helps explain the cross-section of asset returns, would be another fruitful area for future research.

## **8.6. Final Comments and Conclusion**

This Chapter reviewed and summarized the main findings reported by this thesis and their implications for investment management and theoretical asset pricing. The unifying theme of this thesis is the close relation between volatility and asset pricing, both from a time series and from a cross-sectional perspective. In the cross-section of average returns, assets with negative coskewness, and therefore with exposure to volatility risk, command a risk premium on top of the reward for market risk. This relation, as shown in Chapter 4, is empirically strong and bears puzzling implications for the shape of the stochastic discount factor, and thus for the possibility that prices are set by a representative investor. A deeper understanding of the relation between financial volatility, in its widest sense, and asset pricing contributes to the formulation of a richer investment advice and more meaningful policy recommendations. From this perspective and in a context of imperfect (albeit relatively well functioning) and incomplete financial markets, the reduced form statistical representations of volatilities and correlations studied in Chapter 5, 6 and 7 are useful because they provide a better picture of the salient features of the multivariate distribution of asset returns. Thus, in portfolio diversification, portfolio optimization and risk management applications, they help control the relation between first and higher moments of investors' portfolios. In the time series of asset and aggregate returns, variation in volatility drives variation in risk premia. The literature on this relation has been discussed in Chapter 2, while original analytical results on the relation between idiosyncratic volatility, aggregate correlation and aggregate return, and between unconditional second moments and the time variation

of conditional first and second moments were presented in Chapter 5 and 7, respectively.



### 8.6.1 Appendix A: Third Order Utility Expansion

Consider the local variation in a utility function defined over wealth,  $u = u(W_t)$ , given by a third order Taylor expansion:

$$du(W) \cong u'(W)dW + \frac{1}{2}u''(W)dW^2 + \frac{1}{6}u'''(W)dW^3 \quad (\text{A.1})$$

Thus,

$$u'(W) \cong \frac{du(W)}{dW} \cong u'(W) + \frac{1}{2}u''(W)dW + \frac{1}{6}u'''(W)dW^2 \quad (\text{A.2})$$

In discrete time, the local variation of utility is:

$$\Delta u(W_{t+1}) \cong u'(W_t)\Delta W_{t+1} + \frac{1}{2}u''(W_t)\Delta W_{t+1}^2 + \frac{1}{6}u'''(W_t)\Delta W_{t+1}^3 \quad (\text{A.3})$$

The corresponding third order expansion of utility is:

$$\begin{aligned} u(W_{t+1}) &\cong u(W_t) + u'(W_t)\Delta W_{t+1} + \frac{1}{2}u''(W_t)\Delta W_{t+1}^2 + \frac{1}{6}u'''(W_t)\Delta W_{t+1}^3 \\ &= u(W_t) + u'(W_t)W_t \frac{\Delta W_{t+1}}{W_t} + \frac{1}{2}u''(W_t)W_t^2 \frac{\Delta W_{t+1}^2}{W_t^2} + \frac{1}{6}u'''(W_t)W_t^3 \frac{\Delta W_{t+1}^3}{W_t^3} \quad (\text{A.4}) \\ &= u(W_t) + u'(W_t)W_t R_{t+1} + \frac{1}{2}u''(W_t)W_t^2 R_{t+1}^2 + \frac{1}{6}u'''(W_t)W_t^3 R_{t+1}^3 \end{aligned}$$

And marginal utility growth is:

$$\begin{aligned} \frac{u'(W_{t+1})}{u'(W_t)} &\cong \frac{u'(W_t) + u''(W_t)\Delta W_{t+1} + \frac{1}{2}u'''(W_t)\Delta W_{t+1}^2}{u'(W_t)} \\ &\cong 1 + \frac{u''(W_t)}{u'(W_t)} \Delta W_{t+1} + \frac{1}{2} \frac{u'''(W_t)}{u'(W_t)} \Delta W_{t+1}^2 \\ &\cong 1 + \frac{u''(W_t)}{u'(W_t)} W_t \frac{\Delta W_{t+1}}{W_t} + \frac{1}{2} \frac{u'''(W_t)}{u'(W_t)} W_t^2 \frac{\Delta W_{t+1}^2}{W_t^2} \\ &\cong 1 + \frac{u''(W_t)}{u'(W_t)} W_t R_{t+1} + \frac{1}{2} \frac{u'''(W_t)}{u'(W_t)} W_t^2 R_{t+1}^2 \end{aligned} \quad (\text{A.5})$$

Here,  $R_{t+1} = \frac{W_{t+1} - W_t}{W_t} = \frac{\Delta W_{t+1}}{W_t}$ . Then, the first order conditions (2.1) for the maximization of expected utility imply, by identification with (2.8),  $a_t = 1$ ,  $b_{1t} = \frac{u''(W_t)}{u'(W_t)} W_t$  and  $b_{2t} = \frac{1}{2} \frac{u'''(W_t)}{u'(W_t)} W_t^2$ . The quantity  $RRR_t = -\frac{u''(W_t)}{u'(W_t)} W_t$  in the  $b_{1t}$  parameter is the Pratt-Arrow relative risk aversion coefficient, e.g. Pratt (1964). Therefore,  $b_{1t} = -RRR_t$ . The additional parameter  $b_{2t}$  is related to the notion of absolute prudence, defined by Kimball (1990) as  $-\frac{u'''(W_t)}{u''(W_t)}$ . Negative prudence (positive  $b_{2t}$ ) gives rise to preference for portfolio skewness and for a precautionary saving motive in the face of future uncertainty.

Since utility functions are equivalent up to a linear transformation, we might let  $u(W_t) = 0$  and  $u'(W_t) = 1$ . This standardization is often very useful when working with utility functions in that it simplifies their manipulation. From (A.4), the third order Taylor expansion of this standardized utility function around an initial level of wealth  $W_t = 1$  is therefore:

$$\begin{aligned} u(W_{t+1}) &\cong R_{t+1} + \frac{1}{2} u''(1) R_{t+1}^2 + \frac{1}{6} u'''(1) R_{t+1}^3 \\ &\cong R_{t+1} + \theta_1 R_{t+1}^2 + \theta_2 R_{t+1}^3 \end{aligned} \tag{A.6}$$

Here,  $\theta_1 = \frac{1}{2} u''(1)$  and  $\theta_2 = \frac{1}{6} u'''(1)$ . Marginal utility and marginal utility growth then become:

$$u'(W_{t+1}) = 1 + u''(1) R_{t+1} + \frac{1}{2} u'''(1) R_{t+1}^2 = 1 + 2\theta_1 R_{t+1} + 3\theta_2 R_{t+1}^2 \tag{A.7}$$



$$\frac{u'(W_{t+1})}{u'(W_t)} = u'(W_{t+1}) \cong 1 + 2\theta_1 R_{t+1} + 3\theta_2 R_{t+1}^2 \quad (\text{A.8})$$

Letting  $W_t = E_T(W_t) = 1$  yields an expansion around the sample average of wealth or, equivalently, around the sample average of the return on wealth:

$$\begin{aligned} u(W_{t+1}) &\cong \frac{W_{t+1} - E_T(W_t)}{E_T(W_t)} + \theta_1 \left[ \frac{W_{t+1} - E_T(W_t)}{E_T(W_t)} \right]^2 + \theta_2 \left[ \frac{W_{t+1} - E_T(W_t)}{E_T(W_t)} \right]^3 \\ &\cong [R_{t+1} - E_T(R_t)] + \theta_1 [R_{t+1} - E_T(R_t)]^2 + \theta_2 [R_{t+1} - E_T(R_t)]^3 \end{aligned} \quad (\text{A.9})$$

Marginal utility and marginal utility growth can be derived as above.

### 8.6.2 Appendix B: The Pratt-Arrow Risk Premium

In this Appendix, I provide a derivation of the equilibrium relation between expected variance and expected return. Define the simple gamble (an actuarially neutral gamble) as follows:

$$Z \sim \Phi(0, \sigma_z^2) \tag{B.1}$$

With:

$Z$  = a random number (the gamble)

$\Phi(0, \sigma_z^2)$  = a probability distribution with zero mean and  $\sigma_z^2$  variance.

I then assume that investors' utility is a function of wealth ( $W$ ) only:

$$u = u(W) \tag{B.2}$$

Now we can define the condition for the investor to accept the gamble according to the following equation:

$$E[u(W + Z)] = u[W + E(Z) - \pi(W, Z)] \tag{B.3}$$

In equation (B.3) the expression  $\pi(W, Z)$  represents the risk premium that makes the investor indifferent between accepting the actuarially neutral gamble ( $Z$ ) and not accepting. It is assumed to be a function solely of wealth and of the gamble itself.

Now, using equation (B.1), we can substitute out  $E(Z)$  and write:

$$E[u(W + Z)] = u[W - \pi(W, Z)] \tag{B.4}$$

Writing out the Taylor expansion of the left-hand side and right-hand side of equation (B.4):



$$E[u(W + Z)] = E[u(W) + Z u_w + 0.5 Z^2 u_{ww} + \dots] \quad (\text{B.5})$$

$$u[W - \pi(W, Z)] = u(W) - \pi u_w - \frac{1}{2} \pi^2 u_{ww} + \dots \quad (\text{B.6})$$

Here, the terms  $u_w$  and  $u_{ww}$  are the first and second total derivatives of the utility function. Now, equating and simplifying we get:

$$\frac{1}{2} E(Z^2) u_{ww} = -\pi u_w - \frac{1}{2} \pi^2 u_{ww} \quad (\text{B.7})$$

If we further assume the term  $\frac{1}{2} \pi^2 u_{ww}$  to be negligible, we can solve for  $\pi$ , the risk premium:

$$\pi = -\frac{1}{2} \sigma_z^2 u_{ww} / u_w \quad (\text{B.8})$$

The above expression  $-u_{ww}/u_w$  is analogous to the Pratt-Arrow absolute risk-aversion (*ARA*) coefficient. Therefore, using equation (B.8) we can write:

$$\pi = \frac{1}{2} \sigma_z^2 ARA \quad (\text{B.9})$$

Denoting by  $R = (W + Z - \pi - W)/W = (Z + \pi)/W$  the return on the gamble given the risk premium, we can write:

$$\sigma_z^2 = E(Z^2) = E\{W^2 [R - E(R)]^2\} = W^2 \text{Var}(R) \quad (\text{B.10})$$

Therefore, using (B.10) in (B.9) we can write:

$$\pi = \frac{1}{2} W^2 \text{Var}(R) ARA \quad (\text{B.11})$$

Then,

$$-E(R) = \frac{\pi}{W} = \frac{1}{2} W \text{Var}(R) ARA \quad (\text{B.12})$$

$$= \frac{1}{2} RRA \text{Var}(R_\pi) \quad (\text{B.13})$$

Where:

$$RRA = ARA W = -W u_{ww}/u_w$$

Here, the quantity  $\frac{\pi}{W}$  denotes negative expected return that the investor is willing to accept to remove the risk of an otherwise actuarially neutral gamble. The coefficient  $RRA$  in equation (B.13) denotes relative risk aversion. As long as the ‘local shape’ of the utility function does not change, it should be constant against changes in wealth. Equation (B.13) displays a linear relation between risk and expected excess-return that is valid only locally since the relation has been derived on the basis of a second order Taylor expansion of a potentially non-linear equation.



### 8.6.3 Appendix C: Portfolio Diversification

Let  $E_t(\varepsilon_{i,t+1}^2) = \sigma_{i,t}^2 = \text{Var}(R_i)$  represent the variance of the return on an asset  $i$  and  $E_t(\varepsilon_{i,t+1}\varepsilon_{j,t+1}) = \sigma_{ij,t}^2 = \text{Cov}(R_i, R_j)$  represent the covariance between the returns on  $i$  and  $j$ . Neglect possible conditional time variation and multi-period considerations. Consider the variance of portfolio returns:

$$\text{Var}(R_p) = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij} \quad (\text{C.1})$$

Here  $R_p$  indicates portfolio returns,  $N$  is the number of assets in the portfolio,  $w_i$  is the “ $i$ th” asset weight and  $\sigma_{ij}$  is the covariance between the “ $i$ th” and “ $j$ th” asset returns,  $R_i$  and  $R_j$  respectively. The first derivative of  $\text{Var}(R_p)$  with respect to “ $w_i$ ” represents the contribution of the “ $i$ th” asset to portfolio risk:

$$\frac{\partial \text{Var}(R_p)}{\partial w_i} = 2w_i \sigma_{ii}^2 + 2 \sum_{j=1}^N w_j \sigma_{ij} \quad (\text{C.2})$$

Here, the first term  $2w_i \sigma_{ii}^2$  represents the asset-specific (firm-level in an equity portfolio) or idiosyncratic variance contribution and the second term,  $2 \sum_{j=1}^N w_j \sigma_{ij}$ , is the covariance contribution. Suppose that our portfolio strategy consists of a random selection of  $N$  equally weighted assets. Then  $w_i = 1/N \forall i$ . As the number of assets  $N$  increases, the firm-level variance contribution term tends to zero while  $\sum_{j=1}^N w_j$  in the second term, the covariance contribution, tends to unity. Therefore, denoting by  $\bar{\sigma}_{ij} = \sum_{j=1}^N w_j \sigma_{ij}$  the average covariance of asset  $i$  with the other assets, we can write:

$$\lim_{N \rightarrow +\infty} \frac{\partial \text{Var}(R_p)}{\partial w_i} = \lim_{N \rightarrow +\infty} \left( 2w_i \sigma_{ii}^2 + 2 \sum_{j=1}^N w_j \sigma_{ij} \right)$$

$$= \lim_{N \rightarrow +\infty} \left( 2 \sum_{j=1}^N w_j \sigma_{ij} \right) = 2 \bar{\sigma}_{ij} \quad (\text{C.3})$$

The above equation shows that diversification asymptotically eliminates the idiosyncratic, asset-specific variance but not the average covariance contribution to portfolio variance of the constituent assets. This result also obtains in the more general case of a well diversified portfolio with non equal weights, as the latter become on average smaller and smaller as the number of assets increases. The only risk component left in the portfolio is the covariance risk (systematic or un-diversifiable risk), which is the systematic fraction of the total variance of the average asset. Therefore we can rewrite the variance of a highly diversified portfolio as *average covariance*:

$$\text{Var}(R_p) = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij} = \sum_{i=1}^N w_i \sigma_{ip} \quad (\text{C.4})$$

To sum up, using results from Markowitz (1959) mean-variance portfolio theory, I have described two main components of total variance: idiosyncratic variance and systematic variance. I have also shown that the former can be diversified away. The smaller the idiosyncratic variance term, the quicker it will converge to zero as  $N$  increases and  $w_i$  decreases. Therefore, the smaller the idiosyncratic variance of the typical asset, the smaller the number of assets needed to achieve a relatively complete diversification, given a random portfolio selection strategy. On the other hand, the higher the covariance among stock returns, the higher the covariance between the typical stock and the portfolio, i.e. the higher the typical  $\sigma_{ip}$ , and the lower the benefit from diversification because this last term cannot be diversified away. Therefore, the extent to which aggregate portfolio risk can be reduced by diversification is related to the importance of the average covariance term. These well known results have both portfolio management and equilibrium asset pricing implications.



Moments of odd order capture the degree of asymmetry of a distribution. The main third order moment is skewness. Portfolio skewness  $Skew_{p,t} = E_t(\varepsilon_{p,t+1}^3)$  contains both asset skewness  $Skew_{i,t} = E_t(\varepsilon_{i,t+1}^3)$  and coskewness  $Coskew_{ip,t} = E_t(\varepsilon_{i,t+1}\varepsilon_{p,t+1}^2)$  terms<sup>71</sup>. In a portfolio with a large number of assets, the skewness terms are diversified away but not the coskewness terms. The latter represents the non-diversifiable portion of asset skewness. Assets with positive coskewness increase (make more positive) portfolio skewness. Intuitively, assets with positive skewness display positive covariance between their return and the portfolio variance. Thus, adding these assets increases the average covariance between asset returns and portfolio variance and renders the portfolio distribution more skewed to the right.

---

<sup>71</sup> Recall that the terms  $\varepsilon_{i,t}$  and  $\varepsilon_{p,t}$  are zero-mean innovations. Thus, just like variance is defined as a centered second moment and covariance as a centered cross-second moment, skewness is defined as a centered third moments and coskewness as centered cross-third moments

#### 8.6.4 Appendix D: Average Correlation

Proposition: the average correlation between the stocks of a well diversified portfolio can be approximated by the ratio of portfolio variance to the square of the average volatility of the constituent stocks.

Proof:

$$\begin{aligned}\sigma_{p,t}^2 &= \sum_{i=1}^N \sum_{j=1}^N w_{i,t} w_{j,t} \sigma_{i,t} \sigma_{j,t} c_{ij,t} \\ &= \sum_{i=1}^N w_{i,t}^2 \sigma_{i,t}^2 + CORR_t \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N w_{i,t} w_{j,t} \sigma_{i,t} \sigma_{j,t}\end{aligned}\tag{D.1}$$

Here,  $R_{p,t}$  denotes the return on portfolio  $P$ ,  $R_{i,t}$  is the return on the  $i$ th asset,  $N$  is the number of assets included in the portfolio,  $\sigma_{p,t}^2 = Var(R_p)$  is the portfolio variance,  $\sigma_{i,t}^2 = Var(R_{i,t})$  is the variance of asset  $i$ ,  $c_{ij,t}$  is the correlation between asset  $i$  and  $j$ ,  $CORR_t$  is the average asset correlation, i.e. the level of correlation that, if assumed to hold for all pairs of assets, would give the same portfolio volatility as the full correlation matrix.

For ease of algebraic manipulation and to facilitate intuition, it is convenient to define  $\sigma_{ind,t}^2 = \sum_{i=1}^N w_{i,t}^2 \sigma_{i,t}^2$  as the variance that the portfolio would exhibit if all the assets were independently distributed and  $\sigma_{perf,t}^2 = \sum_{i=1}^N \sum_{j=1}^N w_{i,t} w_{j,t} \sigma_{i,t} \sigma_{j,t}$  as the portfolio variance if all the assets were perfectly correlated. Then we can rewrite (D.1) as follows:

$$\sigma_{p,t}^2 = \sigma_{ind,t}^2 + CORR_t (\sigma_{perf,t}^2 - \sigma_{ind,t}^2)\tag{D.2}$$

Finally, solving (D.2) for  $CORR_t$ ,



$$CORR_t = \frac{\sigma_{p,t}^2 - \sigma_{ind,t}^2}{\sigma_{perf,t}^2 - \sigma_{ind,t}^2} \tag{D.3}$$

$$\xrightarrow{N} \frac{\sigma_{p,t}^2}{\sigma_{perf,t}^2}$$

The last step in (D.3) holds asymptotically because for a well diversified portfolio  $\sigma_{ind,t}^2 \xrightarrow{N} 0$ . Define  $\bar{\sigma}_t = \sum_{i=1}^N w_{i,t} \sigma_{i,t}$  as the average total volatility of the assets included in the portfolio. Then, since the variance of a portfolio of perfectly correlated assets is equal to the square of the average volatility of the constituent assets,  $\sigma_{perf,t}^2 = \bar{\sigma}_t^2$  and (D.3) can be rewritten as follows:

$$CORR_t \xrightarrow{N} \frac{\sigma_{p,t}^2}{\bar{\sigma}_t^2} = \frac{\text{Systematic Variance}}{(\text{Average Total Volatility})^2} \tag{D.4}$$

The expression in (D.4) thus provides an asymptotically valid measure of average stock correlation that applies to well diversified portfolios. It is also very similar to an estimator of average correlation used by RiskMetrics™ and discussed by Finger (2000).

### 8.6.5 *Appendix E: A Simple Dynamic Strategy Based on Average Correlation*

As an example of an innovative trading strategy based on exposure to the time-varying level of aggregate correlation, suppose that a derivatives trader uses the methodology proposed for estimating the average correlation of a large set of assets and plots its time series as in Figure 7.1. After visually inspecting the series, he concludes that, since average correlation is close to an all time low and value-weighted correlation is also below its historical average (but it seems to be picking up), it would be desirable to take a position that gained from an average correlation increase. Is it possible to construct such a position using commonly traded financial instruments? It turns out that it is possible and it is also relatively easy. Implied average correlation can be traded by trading options on a basket against a basket of options. To buy implied average correlation, e.g. on the Eurostoxx50 index, the investor should buy an ‘at the money’ option on the index and sell ‘at the money’ options on the single stocks proportionally to the weights of the latter in the index. This position can be replicated by delta-hedging its mirror image (this way the investor can trade implied vs. realized average correlation). This is useful to hedge large portfolios of derivatives priced on the basis of given levels of implied average correlations. Risk premia for this trading activity are likely to be high, at least at the beginning, as the market for correlation risk is relatively untapped (trading it would make the market more complete). Estimating the risk premium that would accrue to a market maker that initiated such a trading activity would be a challenging research possibility.



## Bibliography

Akgiray, V. (1989), "Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts", Journal of Business 62, 55-80.

Alexander, C. (2001), "Orthogonal GARCH" in C. Alexander (editor) Mastering Risk, Vol. 2. Financial Times – Prentice Hall.

Alexander, C. (2002), "Principal component models for generating large GARCH covariance matrices" Review of Banking, Finance and Monetary Economics, 1-23.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and H. Ebens (2001), "The Distribution of Stock Return Volatility", Journal of Financial Economics 61, 43-76.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and J. Wu (2004), "Realized Beta: Persistence and Predictability", PIER Working Paper No. 04-018.

Ang, A. and G. Bekaert (1999), "International Asset Allocation with Time-Varying Correlations," NBER working paper No. 7056.

Ang, A. and G. Bekaert (2002), "International Asset Allocation with Regime Shifts", Review of Financial Studies 15, 1137-1187.

Ang, A. and J. Chen (2000), "Asymmetric Correlation of Equity Portfolios", Working Paper, Columbia University.

Ang, A. and J. Chen (2002), "Asymmetric Correlations of Equity Portfolios", Journal of Financial Economics 63, 443-494.

Arditti, F.D. (1967), "Risk and the Required Return on Equity", Journal of Finance 22, 19-36.

Arrow, K. J. (1971), Essays in the Theory of Risk Bearing, Markham Publishing, Chicago.

Attanasio, O and S. Wadhvani (1990), "Does the CAPM Explain Why the Dividend Yield Helps Explain Returns?", London School of Economics Financial Markets Group Discussion Paper no 4.

Audrino, F. and G. Barone-Adesi (2003), "Semiparametric Multivariate GARCH Models for Volatility Asymmetries and Dynamic Correlations", University of Southern Switzerland Working Paper.

Bae, K.-H., Karolyi, A. and R. M. Stulz (2003), "A New Approach to Measuring Financial Contagion", forthcoming in the Review of Financial Studies.

Baele, L. (2002), "Volatility Spillover Effects in European Equity Markets: Evidence from a Regime-Switching Model", Working Paper.

Baig, T. and I. Goldfajn (1999), "Financial Market Contagion in the Asian Crisis", IMF Staff Papers 46(2).

Baks, K.P., Metrick, A. and J. Wachter (2001), "Should Investors Avoid All Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation", Journal of Finance 56, 45-85.

Balke, N.S. and M.E. Wohar (2001), "Explaining stock price movements: is there a case for fundamentals?", Federal Reserve Bank of Dallas Review (Third Quarter) 22-34.

Barber, B.M. and T. Odean (2000), "Trading Is Hazardous To Your Wealth: the Common Stock Investment Performance of Individual Investors", Journal of Finance 55, 773-806.

Barberis, N. and R.H. Thaler (2003), "A Survey of Behavioural Finance", in Handbook of Economics of Finance, Eds. Constantinides, M.H. and R. Stulz, Elsevier Science B.V.

Bauwens L., Laurent, S. and J.V.K. Rombouts (2003), "Multivariate GARCH Models: A Survey", CORE Discussion Paper.

Bekaert, G. and G. Wu (2000), "Asymmetric Volatility and Risk in Equity Markets", Review of Financial Studies 13, 1-42.

Benartzi, S. and R.H. Thaler (2001), "Naïve Diversification Strategies in Defined Contribution Saving Plan", American Economic Review 91, 79-98.

Bera, A., Garcia, P. and J. Roh (1997), "Estimation of Time-Varying Hedging Ratios for Corns and Soybeans: BGARCH and Random Coefficients Approaches", Sankhya 59, 346-368.

Berndt, E.K., Hall, B.H., Hall, R.E. and R.A. Hausman (1974), "Estimation and Inference in Non-linear Structural Models", Annals of Economic and Social Measurement 3, 653-665.

Billio, M., Caporin, M. and M. Gobbo (2003), "Block Dynamic Conditional Correlations Multivariate GARCH Models", Working Paper no 03.03, GRETA, Università Ca' Foscari Venice.



Black, F. (1976), "Studies of Stock Market Volatility Changes", Proceedings of the American Statistical Association, Business and Economics Statistics Section, 177-181.

Black, F., Jensen, M.C. and M. Scholes (1972), "The Capital Asset Pricing Model: Some Empirical Tests," in Studies in the Theory of Capital Markets, Michael C. Jensen, ed. Praeger, New York, 79-121.

Blake, D. and A. Timmerman (1998), "Mutual Fund Performance: Evidence from the UK", European Finance Review 2, 57-77.

Blume, M. (1970), "Portfolio Theory: A Step towards Its Practical Application." Journal of Business 43, 152-174.

Blume, M. and I. Friend (1973), "A New Look at the Capital Asset Pricing Model", Journal of Finance 28, 19-33.

Bloomfield, T., Leftwich, R. and J.B. Long Jr. (1977), "Portfolio Strategies and Performance", Journal of Financial Economics 5, 2001-218.

Bollerslev, T. (1986), "A Generalised Autoregressive Conditional Heteroskedasticity", Journal of Econometrics 31, 307-327.

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

Bollerslev, T. (1990), "Modelling the Coherence in Short-Run Nominal Exchange Rates: a Multivariate Generalised ARCH Approach", Review of Economics and Statistics 72, 498-505.

Bollerslev, T., Chou, R. and K. Kroner (1992), "ARCH Modelling in Finance: a Review of the Theory and Empirical Evidence", Journal of Econometrics 52, 5-59.

Bollerslev, T., Engle R.F. and Nelson, D.B. (1994), "ARCH Models", in R.F. Engle and D.L. McFadden, ed. Handbook of Econometrics, Vol. 4 (Elsevier Sciences).

Bollerslev, T., Engle, R.F. and J.M. Wooldridge (1988), "A Capital Asset Pricing Model with Time Varying Covariances", Journal of Political Economy 96, 116-131.

Bollerslev, T. and J.M. Wooldridge (1992), "Quasi Maximum Likelihood Estimation in Dynamic Models with Time-Varying Covariances", Econometric Reviews 11, 143-172.

Box, G. and G. Jenkins (1970), Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.

Brailsford, T.J. and R.W. Faff (1996), "An Evaluation of Volatility Forecasting Techniques", Journal of Banking and Finance 20, 419-38.

Brock, W.A., Dechert, D., Scheinkman, H. and B. LeBaron (1996), "A Test for Independence Based on the Correlation Dimension", Econometric Reviews 15, 197-235.

Brooks, C., Henry, O.T. and G. Persaud (2002), "Optimal Hedging and the Value of News", Journal of Business 75, 3330-3352.

Campbell, J.Y (1987), "Stock Returns and the Term Structure", Journal of Financial Economics 18, 373-399.

Campbell, J.Y. (1991), "A Variance Decomposition of Stock Returns", Economic Journal 101, 157-179.

Campbell, J.Y (1999), "Asset Prices, Consumption and the Business Cycle", in Taylor, J and Woodford, M. ed. Handbook of Macroeconomics Vol. 1, North-Holland, Amsterdam, 1231-1303.

Campbell, J.Y and L. Hentschel, (1992), "No News is Good News: an Asymmetric Model of Changing Volatility in Stock Returns", Journal of Financial Economics 31, 281-318.

Campbell, J.Y, Lettau, M., Malkiel, B.G. and Y. XU (2001), "Have Individual Stocks Become More Volatile? An Empirical Exploration Of Idiosyncratic Risk", Journal of Finance 56, 1-43.

Campbell, J.Y, Lo, A.W. and A.C. MacKinlay (1997), The Econometrics of Financial Markets, Princeton University Press, Princeton, NJ.

Campbell, J.Y. and R.J. Shiller (1987), "Cointegration and tests of present value models, Journal of Political Economy 95, 1062-88.

Campbell, J.Y. and R.J. Shiller (1988), "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors", Review of Financial Studies 1, 195-228.

Campbell, J.Y and J.Y. Taksler (2003), "Equity Volatility and Corporate Bond Yields", Journal of Finance 58, 2321-2349.

Campbell, R., Koedijk, K. and P. Kofman (2002), "Increased Correlation in Bear Markets: A Downside Risk Perspective", Financial Analysts Journal 58, 87-94

Cappiello, L., Engle, R.F. and K. Sheppard (2003), "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns", ECB Working Paper No. 204.



Chen, N.F., Roll, R. and S.A. Ross (1986), "Economic Forces and the Stock Market", Journal of Business 59, 383-403.

Cheung, Y.W. and L.K. Ng (1992), "Stock Price Dynamics and Firm Size: an Empirical Investigation", Journal of Finance 47, 1985-1997.

Christie, A. (1982), "The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects", Journal of Financial Economics 3, 407-432.

Chou, R.Y. (1988), "Volatility Persistence and Stock Valuation: Some Empirical Evidence Using GARCH", Journal of Applied Econometrics 3, 279-294.

Cochrane, J.H. (1999), "New Facts in Finance", Economic Perspectives Federal Reserve Bank of Chicago 23, 36-58.

Cochrane, J.H. (1999), "Portfolio Advice in a Multifactor World", Economic Perspectives Federal Reserve Bank of Chicago 23, 59-78 (also NBER working paper No. 7170).

Cochrane, J.H. (2001), Asset Pricing, Princeton University Press, Princeton, NJ.

Cochrane, J.H. (2005), Asset Pricing, 2<sup>nd</sup> Ed., Princeton University Press, Princeton, NJ.

Cochrane, J.H. and J. Raà-Requejo (2000), "Beyond Arbitrage: Good Deal Asset Price Bounds in Incomplete Markets", Journal of Political Economy 108, 79-119.

Constantinides, M.G. (2002), "Rational Asset Prices", Journal of Finance 57, 1567-1591.

Constantinides, M.G. and D. Duffie (1996), "Asset Pricing with Heterogeneous Consumers", Journal of Political Economy 104, 219-240.

Copeland, T.E. and J.F. Weston (1991), Financial Theory and Corporate Policy, Addison-Wesley, New York.

Coval, J.D and T.J. Moskowitz (2001), "The Geography of Investment: Informed Trading and Asset Prices", Journal of Political Economy 109, 811-841.

DeLong, J.B., Shleifer, A., Summers, L. and R. Waldmann (1990a), "Noise Trader Risk in Financial Markets", Journal of Political Economy 98, 703-738.

DeLong, J.B., Shleifer, A., Summers, L. and R. Waldmann (1990b), "Positive Feedback Investment Strategies and Destabilizing Rational Speculation", Journal of Finance 45, 375-395.

De Santis, G. and B. Gerard (1997) "International Asset Pricing and Portfolio Diversification with Time-Varying Risk", Journal of Finance 52, 1881-1912.

Diba, B.T. and Grossman, H.I. (1988), "Explosive rational bubbles in stock prices, American Economic Review 78, 520-30.

Dickey, D.A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", Journal of the American Statistical Association 74, 427-31.

Ding, Z., Granger, C. W. J. and R.F. Engle (1993), "A long memory property of stock market returns and a new model", Journal of Empirical Finance 1, 83-106.

Dittmar, R.F. (2002), "Non Linear pricing Kernels, Kurtosis Preference, and Evidence from the Cross-Section of Equity Returns", Journal of Finance 57, 369-402.

Dueker, M.J. (1997), "Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility", Journal of Business and Economic Statistics 15, 26-34.

Durbin, J. (1969), "Tests for Serial Correlation in Regression Analysis Based on the Periodogram of Least-Squares Residuals", Biometrika 56, 1-15.

Durbin, J. (1970), "Testing for Serial Correlation in Least Squares Regression when Some of the Regressors are Lagged Dependent Variables", Econometrica 38, 410-421.

Edwards, S. (2000), "Interest Rates, Contagion and Capital Controls", NBER Working Paper 7801.

Edwards, S. and R. Susmel (2000), "Interest Rate Volatility and Contagion in Emerging Markets: Evidence form the 1990s", NBER Working Paper 7813.

Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife", Annals of Statistics 7, 1-26.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK inflation", Econometrica 50, 987-1008.

Engle, R.F. (2002), "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalised Autoregressive Conditional Heteroskedasticity Models", Journal of Business and Economic Statistics 20, 339-350.

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation and Testing", Econometrica 55(2), 251-276.



Engle, R.F., Ito, T. and W.L. Lin (1990), "Meteor Showers or Heat Waves? Heteroscedastic Intra-daily Volatility in the Foreign Exchange Market", Econometrica 58, 525-542.

Engle, R.F. and K.F. Kroner (1995), "Multivariate Simultaneous Generalised ARCH", in Econometric Theory, Vol. 11, Cambridge University Press.

Engle, R.F. Lilien, D.F. and R.P. Robins (1987), "Estimating Time-Varying Risk Premia in the Term Structure: the ARCH-M Model", Econometrica 58, 525-542.

Engle, R.F. and J. Mezrich (1996), "GARCH for Groups", Risk 9, 36-40.

Engle, R.F. and V.K. Ng (1993), "Measuring and Testing the Impact of news on Volatility", Journal of Finance 48, 1749-1777.

Engle, R.F. and A.J. Patton (2001), "What good is a volatility model?", Quantitative Finance 1, 237-245.

Engle, R.F. and K. Sheppard (2001), "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH", Working Paper, University of California, San Diego.

Engle, R.F. and K. Sheppard (2002), "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH", Working Paper, University of California, San Diego.

Erb, C.B., Harvey, C.R. and T.E. Viskanta (1994), "Forecasting International Equity Correlations", Financial Analyst Journal 50, 32-45.

Fama, E. (1970), "Efficient Capital Markets: a Review of Theory and Empirical Work", Journal of Finance 25, 383-417.

Fama, E. (1976), Foundations of Finance, Basic Books, New York.

Fama E. and K. French (1989), "Business Conditions and Expected Returns on Stocks and Bonds", Journal of Financial Economics 25, 23-49

Fama E. and K. French (1992), "The Cross-Section of Expected Stock Returns", Journal of Finance 47, 427-465.

Fama E. and K. French (1993), "Common Risk Factors in the Returns on Stocks and Bonds", Journal Financial Economics 33, 3-56.

Fama E. and K. French (1995), "Size and Book-to-Market Factors in Earnings and Returns", Journal of Finance 50, 131-155.

Fama, E. and J. MacBeth (1973), "Risk, Return, and Equilibrium: Empirical Tests", Journal of Political Economy 81, 607-636.

Fiorentini, G., Sentana, E. and G. Calzolari (2002), "Maximum Likelihood Estimation and Inference in Multivariate Conditionally Heteroskedastic Dynamic Regression Models with Student-t Innovations", IPSC-JRC, CEMFI, Florence University Working Paper.

Forbes, K. J. and R. Rigobon (2002), "No Contagion, Only Interdependence: Measuring Stock Market Co-movement", Journal of Finance 57, 2223-2261.

Franses, P.H. and D. van Dijk (1996), "Forecasting Stock Market Volatility Using Non-Linear GARCH Models", Journal of Forecasting 15, 229-35.

Fratzschler, M. (2002), "Financial Market Integration in Europe: On the Effects of EMU on Stock Markets", International Journal of Finance and Economics 7, 165-193.

French, K.R., Schwert, G.W. and R. Stambaugh (1987), "Expected Stock Returns and Volatility", Journal of Financial Economics 19, 3-29.

Friedman, M. and L. Savage (1948), "The Utility Analysis of Choices Involving Risk", Journal of Political Economy 56, 279-304.

Friend, I. and R. Westerfield (1980), "Co-Skewness and Capital Asset Pricing", Journal of Finance 35, 897-913.

Froot, K and M. Obstfeld (1991), "Intrinsic bubbles: The case of stock prices", American Economic Review 81, 1189-1214.

Gallant, A.R., Rossi, J. and G. Tauchen (1992), "Stock Prices and Volumes", Review of Financial Studies 5, 199-242.

Ghysels, E., Harvey, A. and E. Renault (1996), "Stochastic Volatility", in G.S. Maddala, ed. Handbook of Statistics 14, North-Holland, Amsterdam.

Glosten, L.R., Jagannathan, R. and D.E. Runkle (1993), "On the Relation Between the Expected Value and the Variance of the Nominal Excess Return on Stocks", Journal of Finance 48, 1779-1801.

Goyal, A. and P. Santa-Clara (2001), "Idiosyncratic Risk Matters!", Anderson Graduate School of Management, UCLA, Working Paper.

Goyal, A. and P. Santa-Clara (2003), "Idiosyncratic Risk Matters!", Journal of Finance 58, 975-1008.



Guiso, L., Haliassos, M. and T. Jappelli (2002), "Household Stockholding in Europe: Where Do We Stand and Where Do We Go?", CSEF Working Paper no 88, Salerno University.

Guo, H. and R. Savickas (2003), "Idiosyncratic Volatility, Stock Market Volatility, and Expected Stock Returns", Federal Reserve Bank of St. Louis Working Paper no. 2003-028A.

Jadadeesh, N. and S. Titman (1993), "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency", Journal of Finance 48, 65-91.

King, M.A. and S. Wadhvani (1990), "Transmission of volatility between stock markets", Review of Financial Studies 3, 6-33.

Hafner, C. M. and P.H. Franses (2003), "A Generalized Dynamic Conditional Correlation Model for Many Asset Returns", Econometric Institute Report EI 2003-18, Erasmus University Rotterdam.

Hamao, Y., Masulis, R.W. and V. Ng (1990), "Correlations in Price Changes and Volatility across International Stock Markets", Review of Financial Studies 3, 281-307.

Hamilton, J.D. (1994), "Time Series Analysis", Princeton University Press, Princeton, NJ.

Hamilton, J.D. and R. Susmel (1994), "Autoregressive Conditional Heteroskedasticity and Changes in Regime", Journal of Econometrics 64, 307-333.

Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", Econometrica 50, 1029-1054.

Hansen, L.P. and S.F. Richard (1987), "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Asset Pricing Models", Econometrica 55, 587-614.

Hardouvelis, G., Malliaropulos, D. and R. Priestley (2000), "EMU and European Stock Market Integration", CEPR Discussion Paper.

Harris, D., B.P.M. McCabe and S.J. Leybourne (2002), "Stochastic Cointegration: Estimation and Inference", Journal of Econometrics 111(2), 363-84.

Harrison, J.M. and D.M. Kreps (1979), "Martingales and Arbitrage in Multiperiod Securities Markets", Journal of Economic Theory 20, 381-408.

Hartley, R. and L. Farrell (2002), "Can Expected Utility Theory Explain Gambling?", American Economic Review 92(3), 613-624.

Harvey, C.R. (1989), "Time-varying Conditional Covariances in Tests of Asset Pricing Models", Journal of Financial Economics 24, 289-317.

Harvey, C. R. (1996), "Time Varying World Market Integration", Journal of Finance 50, 403-444.

Harvey, C. R. (2000), "Drivers of Expected Returns in International Markets", Emerging Markets Quarterly (Fall).

Harvey, C. R. and A. Siddique (1999), "Autoregressive Conditional Skewness", Journal of Financial and Quantitative Analysis 34, 465-488.

Harvey, C. R. and A. Siddique (2000), "Conditional Skewness in Asset Pricing Tests", Journal of Finance 55, 1263-1295.

Henry, O.T., Olekalns, N., Shields, K.K. and R. Lakshman (2003), "Time Variation and Asymmetry in the World Price of Covariance Risk and the Implications for International Diversification", Manuscript.

Hentschel, L. (1995), "All in the Family: Nesting Symmetric and Asymmetric GARCH Models", Journal of Financial Economics 39 71-104.

Ingersoll, J.E. Jr. (1987), Theory of Financial Decision Making, Rowman and Littlefield, Baltimore.

Jagannathan, R. and Z. Wang (1996), "The Conditional CAPM and the Cross-Section of Expected Returns", Journal of Finance 51, 3-53.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk", Econometrica 47, 263-291.

Karolyi, A. (1995), "A Multivariate GARCH Model of International Transmissions of Stock Returns and Volatility", Journal of Business and Economic Statistics 13, 11-25.

Kearney, C. (2000), "The Determination and International Transmission of Stock Market Volatility", Global Finance Journal 11, 31-52.

Kearney, C. and A. Patton (2000), "Multivariate GARCH Modelling of Exchange Rate Volatility Transmission in the European Monetary System", The Financial Review 41, 29-48.



Kearney, C. and V. Potì (2003), "Idiosyncratic Risk, Market Risk and Correlation Dynamics in European Equity Markets", EFA (Glasgow, 2003) Annual Conference Paper nr. 744, <http://ssrn.com/abstract=424521>.

Kearney, C. and V. Potì (2005a), "Correlation Dynamics in European Equity Markets", forthcoming in Research in International Business and Finance (Elsevier).

Kearney, C. and V. Potì (2005b), "Have European Stocks Become More Volatile?", EFMA (Milan, 2005) and FMA (Chicago, 2005) Annual Meeting Conference Paper.

Kimball, M. (1990), "Precautionary Savings in the Small and in the Large", Econometrica 58, 53-73.

Koutmos, G. (1996), "Modelling the Dynamic Interdependence of Major European Stock Markets", Journal of Business Finance and Accounting 23, 975-989.

Koutmos, G. and G.G. Booth (1995), "Asymmetric Volatility Transmission in International Stock Markets", Journal of International Money and Finance 14, 747-762.

Koutoulas, G. and L. Kryzanowsky (1996), "Macrofactor Conditional Volatilities, Time-Varying Risk Premia and Stock Return Behaviour", The Financial Review 31, 1, 169-195.

Kraus, A. and R. H. Litzenberger (1976), "Skewness Preference and the Valuation of Risky Assets", Journal of Finance 31, 1085-1100.

Lamont, O. (1998), "Earnings and expected returns", Journal of Finance 53, 1563-1587.

Lettau, M. and S. Ludvigson (2001), "Consumption, Aggregate Wealth and Expected Stock Returns", Journal of Finance 56, 815-849.

Lettau, M. and S. Ludvigson (2001), "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time-Varying", Journal of Political Economy 109, 1238-1287.

Levy, H. (1969), "A Utility Function Depending on the First Three Moments", Journal of Finance (September).

Lien, D. and Y., Tse (2002), "Evaluating the Hedging Performance of the Constant-Correlation GARCH Models", Applied Financial Economics 12, 791-798.

Lintner, J. (1965), "Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets", Review of Economics and Statistics 47, 13-37.

Longin, F. and B. Solnik, (1995), "Is Correlation in International Equity Returns Constant: 1960 – 1990?", Journal of International Money and Finance 14, 3-26.

Longin, F. and B. Solnik, (2001), "Extreme Correlation of International Equity Markets", Journal of Finance 56, 649-676

Loretan, M. and P.C.B. Phillips (1994), "Testing the Covariance Stationarity of Heavy-Tailed Series: an Overview of the Theory with Applications to Several Financial Data Sets", Journal of Empirical Finance 1, 211-248.

Lucas, R.E. Jr. (1978), "Asset Prices in an Exchange Economy", Econometrica 46, 1429-1446.

Malkiel, B.J. (1990), A Random Walk Down Wall Street, Norton, New York.

Malkiel, B.J. and Y. Xu (1997), "Risk and Return Revised", Journal of Portfolio Management 23, 9-14.

Mandelbrot, B. (1963), "The Variation of Certain Speculative Prices", Journal of Business 36, 394-419.

Markowitz, H.M. (1952), "The Utility of Wealth", Journal of Political Economy 60, 151-158.

Markowitz, H.M. (1959), Portfolio Selection, Wiley (reprinted in 2<sup>nd</sup> edition in 1991 by Blackwell with hindsight comments made by the author).

McCabe, B.P.M., Harris, D. and S.J. Leybourne (2002), "Testing for Stochastic Cointegration and Evidence for Present Value Models", unpublished manuscript.

Mehra, R. and E. Prescott (1985), "The Equity Premium Puzzle", Journal of Monetary Economics 15, 145-161

Merton, R.C. (1973), "An Intertemporal Asset Pricing Model", Econometrica 41, 867-888.

Merton, R.C. (1980), "On Estimating the Expected Return on the Market: an Exploratory Investigation", Journal of Financial Economics 8, 323-361.

Merton, R.C. (1987), "A Simple Model of Capital Market Equilibrium with Incomplete Information", Journal of Finance 42(3), 483-510.

Mizon, G. (1995), "Progressive Modelling of Macroeconomic Time Series: the LSE Methodology", in K.D. Hoover (Eds.), Macroeconomics: Developments, Tensions and Prospects pp. 107-170, Boston, Kluwer Academic Publishing.



Morillo, D. and L. Pohlman (2002), "Large Scale Multivariate GARCH Risk Modelling for Long-horizon International Equity Portfolios", Panagora Asset Management Working Paper.

Moskowitz, T.J. and M. Grinblatt (1999), "Do Industries Explain Momentum?", Journal of Finance 54(4), 1249-1290.

Muth, J.F. (1961), "Rational Expectations and the Theory of Price Movements", Econometrica 29(3), 315-335.

Nelson, D. B. (1990), "Stationarity and persistence in the GARCH(1,1) model", Econometric Theory 6, 318-334.

Nelson, D.B (1991), "Conditional Heteroskedasticity in Asset Returns: a New Approach", Econometrica 59, 347-370.

Newy, W.K. and D.K. West (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", Econometrica 55, 703-8.

Officer, R.R. (1973), "The Variability of the Market Factor of New York Stock Exchange", Journal of Business 46, 434-453.

Pagan, A.R. (1984), "Econometric Issues in the Analysis of Regressions with Generated Regressors", International Economic Review 25, 221-247.

Pagan, A.R. (1996), "The Econometrics of Financial Markets", Journal of Applied Econometrics 3, 87-105.

Pagan, A.R. and G.W. Schwert (1990), "Alternative Models for Conditional Stock Volatilities", Journal of Econometrics 45, 267-90.

Pagan, A.R. and A. Ullah (1988), "The Econometric Analysis of Models with Risk Terms", Journal of Empirical Finance 3, 15-102.

Patton, A.J. (2002a), "Applications of Copula Theory in Financial Econometrics", Unpublished Ph.D. dissertation, University of California, San Diego.

Patton, A.J. (2002b), "Skewness, Asymmetric Dependence and Portfolios", Working Paper, University of California, San Diego.

Patton, A.J. (2003), "Skewness, Asymmetric Dependence and Portfolios", Working Paper, University of California, San Diego.

Peirò, A. (1999) "Skewness in Financial Returns", Journal of Banking and Finance 23, 847-862.

Pesaran, M. H. and B. Pesaran (1997), Working with Microfit 4.0, Oxford University Press, Oxford.

Post, T. and H. Levy (2002), "Does Risk Seeking Drive Asset Prices?", Timbergen Institute Discussion Paper no. 2002-070/2, Erasmus University Rotterdam.

Post, T., Levy, H. and P. van Vliet (2003), "Risk Aversion and Skewness Preference", ERIM Report Series, No ERS-2003-009-F&A, <http://ssrn.com/abstract=411660>.

Poti, V. (2003), "The DCC-VARMA Model: a Simple Procedure for Estimating the Dynamic Conditional Correlation Parameters of Multivariate Stochastic Processes", Working Paper, <http://ssrn.com/abstract=448220>.

Poti, V. (2005a), "Discount Factor and Conditional Return Volatility", forthcoming in the Applied Financial Economics Letters.

Poti, V. (2005b), "Coskewness and Conditional Asset Pricing", FMA Annual Meeting Conference Paper, New Orleans 2004.

Poti, V. (2005c), "The Coskewness Puzzle", European Finance Association (EFA) Doctoral Colloquium Paper, Moscow 2005, <http://ssrn.com/abstract=821315>.

Pratt, J.W. (1964), "Risk Aversion in the Small and in the Large with Applications", Econometrica 32, 122-136.

Ramsey, J.B. (1969), "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis", Journal of the Royal Statistical Society 31(2), 350-71.

Richardson, M. and T. Smith (1993), "A Test for Multivariate Normality in Stock Returns", Journal of Business 66, 295-321.

Richter, M. (1960), "Cardinal Utility, Portfolio Selection and Taxation", Review of Economic Studies (April).

Roll, R. (1977), "A Critique of the Asset Pricing Theory's Tests", Journal of Financial Economics (March), 129-176.

Ross, S.A. (1976), "The Arbitrage Theory of Capital Asset Pricing", Journal of Economic Theory (December), 343-362.



Sargent, T. (1993), Bounded Rationality in Macroeconomics, Oxford University Press, Oxford.

Schwert, G.W. (1989), "Why Does Stock Market Volatility Change Over Time?", Journal of Finance 44, 1115-1153.

Scruggs, T.J. (1998), "Resolving the Puzzling Intertemporal Relation between the Market Risk premium and Conditional Market Variance: A Two Factor Approach", Journal of Finance 53, 575-603.

Sephton, P. (1993), "Optimal Hedge Ratios at the Winnipeg Commodity Exchange", Canadian Journal of Economics 26, 175-193.

Sharpe, W.F. (1964), "Capital Asset Prices: a Theory of Market Equilibrium Under Conditions of Risk", Journal of Finance 19, 425-442.

Shefrin, H. (2005), A Behavioral Approach to Asset Pricing, Elsevier Academic Press.

Shefrin, H. and M. Statman (2000), "Behavioral Portfolio Theory", Journal of Financial and Quantitative Analysis 35(2), 127-151.

Shiller, R.J. (1981), "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?", American Economic Review 71, 421-436.

Shleifer, A. and R.W. Vishny (1997), "The Limits of Arbitrage", Journal of Finance 52(1), 35-55.

Simkowitz, M. and W. Beedles (1978), "Diversification in a Three Moment World", Journal of Financial and Quantitative Analysis 13, 927-941.

Sims, C.A. (1980), "Macroeconomics and Reality", Econometrica 48, 1-48.

Singleton, J. C. and J. Wingender (1986), "Skewness Persistence in Common Stock Returns", Journal of Financial and Quantitative Analysis 21, 335-341.

Statman, M. (2002), "Lottery Players/Stock Traders", Financial Analysts Journal (January/February), 14-21.

Stock, J and M. Watson (1988), "Testing for Common Trends", Journal of the American Statistical Association 55, 1097-1107.

Szego. G. (2002), "Measures of Risk", Journal of Banking and Finance 26, 1253-1272.

Theodossiou, P., Kahya, E., Koutmos, G. and A. Christofi (1997), "Volatility Reversion and the Correlation Structure of Returns in Major International Stock Markets", The Financial Review 32, 205-225.

Tsiang, S. (1972), "The Rational of the Mean-Standard Deviation Analysis, Skewness Preference and the Demand for Money", American Economic Review (June).

Tuner, C.M., Startz, R. and C.R. Nelson (1989), "A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market", Journal of Financial Economics 25, 3-22.

Tversky, A. and D. Kahneman (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty", Journal of Risk and Uncertainty 5, 297-323.

Wermers, R. (2000), "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs, and Expenses", Journal of Finance 55, 1655-1695.

White, H. (1980), "A Heteroskedasticity - Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", Econometrica 48, 817-838.

Whitelaw, R. (1994), "Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns", Journal of Finance 49, 515-541.