Do fixed income securities also show asymmetric effects in conditional second moments?

Lorenzo Cappiello*

Executive Summary

The so-called "asymmetric volatility" phenomenon is one of the empirical regularities shown by (conditional) estimates of equity second moments. Typically, volatility increases more after negative than positive return shocks of the same magnitude, and, sometimes, it even falls subsequent to an increase in stock prices. Two explanations have been put forth for this phenomenon: The leverage effect hypothesis, due to Black (1976) and Christie (1982), and the volatility feedback effect proposed by Campbell and Hentschell (1992) and extended by Wu (2000). Surprisingly, whereas there has been a proliferation of conditional econometric models able to capture asymmetry in volatility (see Hentscell, 1995, for a synthesis), there is a lack of conditional econometric specifications able to explicitly model asymmetry in covariances. However, as argued by Kroner and Ng (1998), if expected returns on one asset change because an asymmetric volatility effect occurs, the covariance between returns on that asset and returns on assets which have possibly not experienced such an effect should also change. Kroner and Ng (1998) and Bekaert and Wu (2000) are the only two pieces of research I am aware of where multivariate asymmetric models for conditional variances *as well as* covariances are developed and tested.

The need to take into account the asymmetric effects on conditional second moments has an appealing economic justification. Assume, for instance, that a negative return shock generates more volatility than a positive return innovation of the same magnitude. When a traditional Generalised Autoregressive Conditionally Heteroskedastic (GARCH) process is used to model second moments, the conditional volatility which occurs after a price drop will be underestimated. Similarly, the conditional volatility which follows a price increase will be overestimated. Consequences such as asset mis-pricing and poor in- and out-of-sample forecasts will be, therefore, unavoidable. Accurate conditional (co)variance estimation of equities as well as other typologies of assets, thus, is crucial for portfolio selection, risk management, and pricing of primary and derivative securities.

While the asymmetric phenomenon in variances has been widely explored for individual stocks, equity portfolios, and/or stock market indices, it has probably never been tested for fixed income securities.

The main goal of this research is to consider a portfolio which includes not only equities but also Treasury bills and government bonds, with the purpose of documenting the asymmetry phenomenon for stocks *and* fixed income securities. Moreover, including a variety of liquid investment opportunities makes the portfolio more realistic and increases the power of the test (Cappiello, 1999).

Two multivariate GARCH processes able to capture the asymmetric effects for both conditional variances *and* covariances are developed and tested. The first parametrization, named "generalised Nelson model", is an extension of the univariate

* The Graduate Institute of International Studies - International Economic Section - Pavillon Rigot, 11A, Avenue de la Paix CH-1202 Geneva, Switzerland. Tel.: +41 (22) 734.89.50; fax: +41 (22) 733.30.49; e-mail: cappiel5@hei.unige.ch.

Exponential GARCH of Nelson (1991). As such, it is characterised by the fact that the terms which capture asymmetry assign the same weight to both past positive and negative innovations. The model, though, is discarded from the analysis since it does not fit the data well. The second specification, baptised "generalised GJR model", is built on the univariate asymmetric GARCH of Glosten, Jagannathan, and Runkle (1993) and Zakoian (1994). As such, the components which accommodate asymmetry assign different weights to past positive and negative shocks. Contrary to the former parameterisation, the latter seems to be well-specified. Conditional second moments for equities *as well as* fixed income securities do respond asymmetrically to past positive and/or negative shocks as evidenced by news impact curves and surfaces as well as robust conditional moment tests (Engle and Ng, 1993, and Kroner and Ng, 1998). In particular, negative shocks have a significant larger impact on equity volatility than positive shocks of the same magnitude, whereas the opposite occurs for Treasury bills. As for government bonds, volatility reacts differently to past positive and negative news.

GARCH processes are statistical models. The theoretical framework to which the two asymmetric GARCH processes described above are applied to is the Intertemporal Capital Asset Pricing Model (CAPM) of Merton (1973). The choice of a dynamic rather than a static CAPM is motivated by the fact that multifactor models seem to have supplanted single factor ones, which are usually mis-priced (see Cochrane, 1999, and Cappiello, 1999, for further details on the topic). Typically, single factor models link expected excess returns of an asset to the market risk exposure. In addition to this, multifactor models consider other state variables relevant to security performance, like oil prices, inflation, business cycle proxies, interest rates, etc. Numerous studies have pointed out that returns and volatility of stocks and bonds are linked to the business cycle. Most of this research finds that the stock market falls before (or in concomitance of) economic recessions, anticipating (or matching) a downturn in the business cycle, and rises before (or in concomitance of) troughs, anticipating (or matching) a recovery in the economy (see Cappiello, 1999, for a survey). In the view of this relationship, the growth rate of industrial production is chosen here as a second priced factor. While in the literature one can find several empirical investigations of single-factor multi-asset CAPMs as well as of multi-factor single-asset CAPMs, a conditional estimation of a multi-factor multi-asset CAPM carried out with a GARCH-in-Mean (GARCH-M) technique represents a novelty, attempted, until now, only by Cappiello (1999). The growth rate of industrial production turns out to be significantly priced and relevant in the determination of the total premia required to hold risky securities. Furthermore, major financial market turmoil as well as spillovers from one market to another is reflected by the time evolution of risk premia.

Testing a CAPM implies the estimation of the prices of risk it involves. Such an estimation, though, is one of the "unsolved" issues in empirical finance. Since prices of risk are related to the investors' utility function, which is not observable *per se*, their evolution over time is often based on some presumptions. Nevertheless, their sign and magnitude are crucial, since they directly affect risk premia.

There is large consensus on the fact that both expected excess returns and conditional (co)variances change through time (see the extensive surveys of Bollerslev, Chou, and Kroner, 1992, Bera and Higgins, 1993, and Bollerslev, Engle, and Nelson, 1994). However, first and second conditional moments do not move in a one-to-one proportion. Therefore, in CAPM-type models the risk-return relationship is not constant over time. With few exceptions (see, for instance, De Santis and Gerard, 1998b, and Cappiello, 1999), the literature is silent about the possibility that also prices of intertemporal risk vary over time. However, if it is true that the risk-return relationship changes because both first and second moments vary over time not in a one-to-one proportion, for the same reason intertemporal prices of risk have to be time-varying.

Following Cappiello (1999), the prices of risk, first held constant, are next allowed to vary according to the regime switching model of Hamilton (1988, 1989, 1990, 1994). The use of Hamilton's filter, which represents a novelty in the estimation of prices of risk, is combined only with the generalised GJR model, due to its better performance. Two regimes are identified, one in which the price of market risk is high and the price of intertemporal risk is low, and one in which the reverse occurs, i.e. the first price is high and the second low. This can be interpreted as a switch in investors' preferences whose degree of risk aversion increases in correspondence to or after financial turmoil. However, if the specification where prices of risk can change through time accommodates shifts in agents' preferences, it sacrifices flexibility in terms of GARCH specification. Thus, relevant phenomena, like the decrease in the bond risk premia which occurred after some equity market falls, are not captured, whereas they are when prices are held constant.

Do Fixed Income Securities Also Show Asymmetric Effects in Conditional Second Moments?

Lorenzo Cappiello*

January 2000

Abstract

This paper estimates a trivariate two-factor conditional version of the Intertemporal CAPM of Merton (1973). The three considered assets are: US stocks, 6-month T-bills, and 10-year government bonds. As a second factor the growth rate of industrial production is chosen. Two multivariate GARCH processes able to capture the asymmetric effects for both conditional variances *and* covariances are developed and tested. News impact curves and surfaces as well as robust conditional moment tests (Engle and Ng, 1993, and Kroner and Ng, 1998) indicate that conditional second moments for equities *as well as* fixed income securities do respond asymmetrically to past positive and/or negative news. Finally, the prices of market and intertemporal risk, first held constant, are next allowed to vary over time according to the regime switching model of Hamilton (1988, 1989, 1990, 1994). The two identified states might reflect a switch in investors' preferences whose degree of risk aversion increases in correspondence to or after financial turmoil.

Keywords: Intertemporal CAPM, business cycle, asymmetric multivariate GARCH-in-Mean, regime shifts.

JEL classification: C32; E32; G12.

---

*The Graduate Institute of International Studies, University of Geneva - Pavillon Rigot, 11A, Avenue de la Paix 1202 Geneva, Switzerland. Tel.: +41 (22) 734.89.50; fax: +41 (22) 733.30.49; e-mail: cappiel5@hei.unige.ch. I am grateful to Tom Fearnley, Hans Genberg, Christian Gouriéroux, Campbell Harvey, Arjan Kadareja, Henri Loubergé, Alain Monfort, Federica Sbergami, and Charles Wyplosz for their valuable comments and discussions. A special thank to Urs Luterbacher for letting me use his workstation. Financial support from the International Center for Financial Asset Management and Engineering (FAME) is gratefully acknowledged. Of course, remaining errors are mine alone.

Contents

1 Introduction: Empirical Regularities and Research Goals

The so-called "asymmetric volatility" phenomenon is one of the empirical regularities shown by (conditional) estimates of equity second moments. Typically, volatility increases more after negative than positive return shocks of the same magnitude, and, sometimes, it even falls subsequent to an increase in stock prices. Two explanations have been put forth for this phenomenon: The leverage effect hypothesis, due to Black (1976) and Christie (1982), and the volatility feedback effect proposed by Campbell and Hentschell (1992) and extended by Wu (2000)[1]. Surprisingly, whereas there has been a proliferation of conditional econometric models able to capture asymmetry in volatility (see Hentscell, 1995, for a synthesis), there is a lack of conditional econometric specifications able to explicitly model asymmetry in covariances. However, as argued by Kroner and Ng (1998), if expected returns on one asset change because an asymmetric volatility effect occurs, the covariance between returns on that asset and returns on assets which have possibly not experienced such an effect should also change. Kroner and Ng (1998) and Bekaert and Wu (2000) are the only two pieces of research I am aware of where multivariate asymmetric models for variances *as well as* covariances are developed and tested.

The need to take into account the asymmetric effects on conditional second moments has an appealing economic justification. Assume, for instance, that a negative return shock generates more volatility than a positive return innovation of the same magnitude. When a traditional Generalized Autoregressive Conditionally Heteroskedastic (GARCH) process is used to model second moments, the conditional volatility which occurs after a price drop will be underestimated. Similarly, the conditional volatility which follows a price increase will be overestimated. Consequences such as asset mis-pricing and poor in- and out-of-sample forecasts will be, therefore, unavoidable. Accurate (co)variance estimations of equities as well as other typologies of assets, thus, is crucial for portfolio selection, risk management, and pricing of primary and derivative securities.

While the asymmetric phenomenon in variances has been widely explored for individual stocks, equity portfolios, and/or stock market indices, it has probably never been tested for fixed income securities.

The main goal of this research is to consider a portfolio which includes not only equities but also Treasury bills and government bonds, with the purpose of documenting the asymmetry phenomenon for stocks *and* fixed

[1]Campbell and Hentschell (1992) present for the first time a fully worked out formal model on volatility feedback. However, this hypothesis can also be tracked down in, among others, Malkiel (1979), Pindyck (1984), Poterba and Summers (1986), and French, Schwert, and Stambaugh (1987).

income securities. Moreover, including a variety of liquid investment opportunities makes the portfolio more realistic and increases the power of the test (Cappiello, 1999).

Two multivariate GARCH processes able to capture the asymmetric effects for both conditional variances *and* covariances are developed and tested. The first model extends the well-known univariate Exponential GARCH (EGARCH) of Nelson (1991). As such, it is characterised by the fact that the terms which capture asymmetry assign the same weight to both past positive and negative innovations. One can name this specification "generalized Nelson model". In the spirit of Ding and Engle (1994), it is assumed that the GARCH process is covariance stationary. This hypothesis reduces the number of parameters to estimate and facilitates convergence.

The second specification proposed here is a modified version of the multivariate asymmetric BEKK GARCH model suggested by Kroner and Ng (1998)[2]. This latter parametrization can be viewed as an extension to a multivariate context of the popular asymmetric univariate GARCH process of Glosten, Jagannathan, and Runkle (1993). As such, the components which accommodate asymmetry assign different weights to past positive and negative shocks. Two major changes to Kroner and Ng's (1998) model are suggested. As before, it is assumed that the process is covariance stationary, which permits to economize considerably on the number of parameters to estimate. Moreover, Kroner and Ng's (1998) representation is designed for equity portfolios only. However, Engle and Ng's (1993) tests for asymmetry show that fixed income securities also exhibit asymmetric effects, but of a different type than those of equities (see Section 3 for further details). Therefore, in the multivariate GARCH process developed here, the component which captures asymmetry will account for the special characteristics of each asset. In particular, equities show that negative shocks have a significant larger impact on volatility than positive shocks of the same magnitude, whereas the opposite occurs for T-bills. As for government bonds, volatility seems to react with a different intensity to past positive and negative shocks. One can baptize this GARCH representation "generalized GJR model".

The two GARCH specifications proposed here are estimated and compared. The robust conditional moment tests suggested by Kroner and Ng (1998) are applied to the chosen model to check whether it accommodates the skewness shown by the data. While asymmetry in conditional variances is reflected by the "news impact curves" of Engle and Ng (1993), asymmetry in conditional covariances will be shown by the "news impact surfaces" of

[2]BEKK is an acronym from the initials of the researchers (Baba, Engle, Kraft, and Kroner) who developed a multivariate symmetric GARCH parametrization widely used in the literature (see Engle and Kroner, 1995, for further details). Kroner and Ng (1998) have extended this model and rendered it able to capture asymmetry effects.

Kroner and Ng (1998).

GARCH processes are statistical models. The theoretical framework to which the two asymmetric GARCH processes described above are applied to is the Intertemporal Capital Asset Pricing Model (CAPM) of Merton (1973). The choice of a dynamic, rather than a static, CAPM is motivated by the fact that multifactor models seem to have supplanted single factor ones, which are usually mis-priced (see Cochrane, 1999, and Cappiello, 1999, for further details on the topic). Typically, single factor models link expected excess returns of an asset to the market risk exposure. In addition to this, multifactor models consider other state variables relevant to security performance, like oil prices, inflation, business cycle proxies, interest rates, etc. Numerous studies have pointed out that returns and volatility of stocks and bonds are linked to the business cycle. Most of this research finds that the stock market falls before (or in concomitance of) economic recessions, anticipating (or matching) a downturn in the business cycle, and rises before (or in concomitance of) troughs, anticipating (or matching) a recovery in the economy (see Cappiello, 1999, for a survey). In the view of this relationship, the growth rate of industrial production will be chosen here as a second priced factor. While in the literature one can find several empirical investigations of single-factor multi-asset CAPMs as well as of multi-factor single-asset CAPMs, a conditional estimation of a multi-factor multi-asset CAPM carried out with a GARCH-in-Mean (GARCH-M) technique represents a novelty, attempted, until now, only by Cappiello (1999). Furthermore, following Cappiello (1999), the prices of risk implied by Merton's (1973) asset pricing theory are first kept constant and then allowed to vary according to the regime switching model of Hamilton (1988, 1989, 1990, 1994). The fact that the risk-return relationship is time-varying has been well documented in the literature (see Cappiello, 1999, for a survey), however its precise dynamic is unknown. The use of Hamilton's filter represents a novelty in the estimation of prices of risk.

The paper is organized as follows. In the next sub-Section the two explanations put forth to justify the asymmetric volatility phenomenon are discussed and the literature on the topic is briefly reviewed. A synthesis of Merton's (1973) asset pricing theory is presented in Section 2, whereas Section 3 describes the data employed in the analysis. Section 4 discusses the econometric methodology, while the empirical results are presented in Section 5. Section 6 summarizes the main results obtained and concludes the paper.

1.1 The Asymmetric Volatility Phenomenon

Black (1976) and Christie (1982) were among the first to document that volatility increases more after stock market falls than rises. They explained this regularity through the financial leverage effect. More clearly, since an unexpected drop in a stock value increases the debt-to-equity ratio of a company, the riskiness of that stock will surge as well. The leverage effect alone, however, seems to be too small to account for the whole asymmetric volatility phenomenon (Black, 1976, Christie, 1982, and Schwert, 1989), hence the need for different theories.

Campbell and Hentschell's (1992) explanation is more complicated and is based on a CAPM-type specification of the mean equation. Its validity rests on the existence of time-varying risk premia, volatility persistence, and a positive relationship between expected asset returns and risk exposure. The fact that risk premia are time varying is well documented in the literature, as shown by the extensive surveys of Bollerslev, Chou, and Kroner (1992), Bera and Higgins (1993), and Bollerslev, Engle, and Nelson (1994). Volatility clustering has also been well recognized at least since the early '60s (see, for instance, Mandelbrot, 1963, or Fama, 1965) and largely tested by the huge literature on conditional second moments (see the above mentioned surveys). As for the positive risk-return relationship, the evidence is mixed. Some research finds a significant negative or an insignificant positive relationship, like, for instance, in French, Schwert, and Stambaugh (1987), Campbell (1987), Turner, Startz, and Nelson (1989), Baillie and DeGennaro (1990), Nelson (1991), Glosten, Jagannathan, and Runkle (1993), De Santis and İmrohoğlu (1997), and Bekaert and Wu (2000). Others, like Bollerslev, Engle, and Wooldridge (1988), Harvey (1989), Campbell and Hentschell (1992), Friedman and Kuttner (1992), Bonomo and Garcia (1997), Scruggs (1998), and Cappiello (1999), find, instead, a positive significant relationship. Without entering the issue of the sign of the risk-return relationship and its statistical significance, for which one can refer to Cappiello (1999), it seems that the use of multi-factor multi-asset models improves the efficiency and the power of the test, making the price of market risk positive and significant.

The volatility feedback effect distinguishes between two cases, bad and good news (or shocks). Following Engle and Ng (1993), a piece of news, ε_t, is defined as the difference between the rate of return on an asset, R_t, and its expected value conditional on the information set available at time $t-1$, $\Im_{t-1}$, i.e. $M_t \equiv E\left(R_t \left| \Im_{t-1}\right.\right)$; therefore $\varepsilon_t = R_t - M_t$. A negative shock indicates the arrival of bad news, while a positive shock signals that good news is coming up. Consider, first, the bad news case. Assume that investors expect an increase in future stock volatility. Taking Wu's (2000) example, this might be due to foreign financial market turmoils, which, in turn, can spill over into domestic markets. As a consequence, investors will be willing to sell and reluctant to buy, causing a drop in equity prices. In other words, if volatility is expected to rise and the CAPM holds, i.e. volatility is positively priced, then investors will require a compensation in terms of higher expected returns. The consequent fall in prices and stickiness of volatility feedback causing further return shocks and volatility rises, which, in turn, exacerbate the initial price drop. The process goes on until the expected returns are sufficiently high. Suppose, in contrast, that good news arrives to the market. In other words, assume that investors expect a decrease in volatility. This might again be due to a leading financial market which surges for internal reasons: Prices rise, but the associated risk either goes up less than proportionately or even decreases. Traders, this time, will be willing to buy rather than sell, moving asset prices up, or, equivalently, expected returns down. The positive shock, however, will trigger an increase in volatility. Again, if volatility is positively priced and persistent (which, in turn, generates an upward revision of future volatility), investors will require higher expected returns, with a consequent drop in prices. This price decline, which tends to offset the initial price increase, generates a negative shock and, therefore, once more an increase in volatility. Such an increase in volatility after the arrival of good news tends to counterbalance its initial anticipated decrease. Thus the final effect on asset volatility is uncertain. Definitely, volatility will increase less than in the case of a negative shock of the same magnitude. Finally, assume that no news arrives. Under this extreme case, "the market rises because 'no news is good news' about future volatility." (Campbell and Hentschell, 1992).

As pointed out by Bekaert and Wu (2000), the leverage effect and the feedback volatility hypothesis are not in conflict and can, in fact, be functioning at the same time. When bad news arrives into the market, triggering a decline in stock prices, the leverage effect starts to play its role, reinforcing the volatility feedback effect. On the other hand, if good news arrives, equity prices first go up and then down. When the whole volatility feedback mechanism has unfolded, prices can either increase or decrease, with ambiguous consequences on firms' leverage as well as on the final impact on stock return volatility. More precisely, if prices go up, the debt-to-equity ratio decreases and with it the associated stock risk; if prices go down, the reverse occurs: Firms' leverage deteriorates and the risk becomes higher. Obviously, when assets like T-bills and/or government bonds are considered, the leverage effect does not play any role. Once again, though, the arrival of bad news is expected to increase volatility unambiguously, whereas the arrival of good news should produce an uncertain effect on volatility.

Finally, notice the two theories on asymmetric volatility propose a reversed causality. Whereas under the leverage effect hypothesis an unexpected price drop triggers a volatility surge, under the feedback effect assumption an anticipated change in volatility causes a price decline. Moreover, Campbell and Hentschell's (1992) explanation is more general and can turn out to be appropriate not only for equities but also for other categories of assets, like fixed income securities.

Wu (2000) extends Campbell and Hentschell's (1992) asymmetric volatility model. In Campbell and Hentschell stock returns are related to dividend volatility, rather than stock volatility, as it occurs in the classical CAPM. Furthermore, excess returns also depend positively on dividend shocks and negatively on the square of dividend shocks. The last two variables, which in fact simplify into the same state variable since one is the square of the other, proxy news about dividends and their volatility. The model captures asymmetric conditional variances and accommodates the negative skewness and excess kurtosis of the data. In Wu news about dividends and their volatility are explained by two state variables instead of one. As in Campbell and Hentschell, returns are positively linked to dividend volatility and dividend shocks, where the latter are aimed at capturing the impact of dividend news. The difference is in the variable that describes news about dividend volatility. In Wu this is proxyed by the stochastic component of the variance of the dividend growth rate, which is assumed to follow a stochastic volatility process. Estimates of the model point out that both the leverage and the volatility feedback effects play a role in explaining asymmetric volatility. Finally, note that the rationale for considering dividends lies in the fact that the present value of a firm depends on the stream of its dividends.

Independently of the possible explanations, there exists a vast empirical literature that seeks to capture the asymmetric volatility phenomenon. An incomplete list includes Black (1976), Christie (1982), French, Schwert, and Stambaugh (1987), Engle (1990), Schwert (1990), Nelson (1991), Sentana (1991), Campbell and Hentschell (1992), Cheung and Ng (1992), Gouriéroux and Monfort (1992), Ding, Granger, and Engle (1993), Engle and Ng (1993), Glosten, Jagannathan, and Runkle (1993), Rabemananjara and Zakoian (1993), Hamilton and Susmel (1994), Zakoian (1994), Booth and Koutmos (1995), Braun, Nelson, and Sunier (1995), Duffee (1995), Hentschell (1995), Hamilton and Lin (1996), Koutmos (1996), Bekaert and Harvey (1997), Booth, Martikainen, and Tse (1997), Fornari and Mele (1997), Kroner and Ng (1998), Lubrano (1998), Scruggs (1998), Susmel (1998), Bekaert and Wu (2000), Wu (2000), and Wu and Xiao (2000). Except for Kroner and Ng and Bekaert and Wu, none of these studies model and test for asymmetry in conditional covariances. This is because either only one asset is considered and, therefore, there are no covariances to estimate, or because, even when multi-asset models are analyzed, covariances do not explicitly account for any asymmetric effect.

Kroner and Ng (1998) devise a general multivariate GARCH model which nests four popular GARCH specifications and is able to capture asymmetric effects in variances and covariances[3]. Such a model, when applied to large- and small-firm portfolios, detects significant asymmetric effects in both variances and covariances. Specifically, small-firm return shocks have negligible effects on variances and covariances. However, large-firm negative shocks affect volatility of both small- and large-firm returns. Moreover, bad news about large-firm returns has a higher impact on conditional covariances between small- and large-firm returns than good news.

Bekaert and Wu (2000) develop a model able to disentangle the leverage effect from the volatility feedback hypothesis. Their results are based on a conditional static CAPM estimated with an asymmetric multivariate GARCH specification which is built up in the spirit of the BEKK model suggested by Kroner and Ng. The analysis is applied to four assets: Three portfolios of 5 equities each (pulled from the Nikkei 225 stocks), representing firms with low, medium, and high leverage ratios, respectively, plus a market portfolio. The market, the high and medium leverage portfolios show a strong asymmetric effect mainly due to market shocks. Asymmetry for the low leverage portfolio is more modest and is essentially generated by firm-specific shocks. The asymmetric effect in the market portfolio is uncovered with asymmetry in conditional volatility, whereas for the high, medium, and low leverage portfolios it is the conditional covariance asymmetry which governs the asymmetric effects. In particular, while bad news increases conditional covariances, good news has a mixed impact on them.

2 The Intertemporal Capital Asset Pricing Model

The asset pricing model that drives investors' optimal portfolio choices is assumed to be the conditional version of the Merton's (1973) continuous-time Intertemporal CAPM[4]. The main intuitions behind this theory, the asset pricing restrictions which it implies, and the discussion on the prices of risk are briefly reviewed here. For further details, the reader is referred to Cappiello (1999).

While the static CAPM of Sharpe (1964), Lintner (1965), Mossin (1966), and Black (1972) is derived under the assumption that investors live for only

[3]The four GARCH specifications included as special cases are the diagonal VECH model of Bollerslev, Engle, and Wooldridge (1988), the BEKK model of Engle and Kroner (1995), the Factor ARCH model of Engle, Ng, and Rothschild (1990), and the constant correlation model of Bollerslev (1990).

[4]Long (1974) and Fama (1996) provide discrete-time versions of the model.

one period, in the real world consumption and investment decisions span longer horizons. In such a dynamic economy, the investment opportunity set changes over time; these changes are governed by one or more state variables, X_l, $l = 1, \ldots, m$. Risk-averse rational agents will thus anticipate and hedge against the possibility that investment opportunities may adversely change in the future. Because of this hedging need, the equilibrium expected returns on securities will depend not only on "systematic" or "market" risk (as in the traditional CAPM), but also on "intertemporal" risks. As is well known, market risk is measured by the covariance of asset returns with the market returns. Similarly, intertemporal risks are given by the covariance of security returns with state variables.

In line with Merton (1973), it is assumed that both rates of returns on assets and state variables follow a standard Brownian motion and that the risk-averse representative agent maximizes his expected intertemporal utility function subject to a wealth constraint:

$$\max_{C,w_i} E \int_t^T U\left(C(s), s\right) ds \tag{1}$$

$$\text{s.t. } dW = \left[\sum_{i=1}^n w_i \left(E(r_i) - r_f\right) + r_f\right] W dt - C dt + W \sum_{i=1}^n w_i \sigma_i dz_i,$$

where $C(t)$ is the instantaneous consumption flow, W the wealth value, w_i the fraction of wealth invested in security i, $E(r_i)$ the instantaneous expected return on asset i, r_f the return on the risk-free asset, σ_i the instantaneous standard deviation of asset return i, and z_i follows a standard Wiener process[5].

Let $J\left(W(t), X(t), t\right)$ be the derived utility function of wealth, i.e. $J\left(W, X, t\right) \equiv \max E \int_t^T U\left(C, s\right) ds$. The solution of the optimization problem yields the following $n + 1$ first order conditions:

$$U_c\left(C, t\right) = J_W\left(W, X, t\right), \tag{2}$$

and

$$E\left(R_{i,t+1}|\Im_t\right) = \lambda_{M,t} \sum_{j=1}^n Cov\left(R_{i,t+1}, R_{j,t+1}, |\Im_t\right) w_{j,t} + \sum_{l=1}^m \lambda_{Fl,t} Cov\left(R_{i,t+1}, X_{l,t+1}, |\Im_t\right), \tag{3}$$

[5]Note that in Merton's (1973) Intertemporal CAPM, r_i, r_f and σ_i, $\forall i$, are time-varying. As such, first and second moments are conditional on the information available up to the time where expectations are formed. Here, however, to simplify the notation, the time dependence of asset returns as well as that of first and second moments is omitted.

for $i = 1, \ldots, n$. Equation (2) is the familiar intertemporal envelope condition, which equates the marginal utility of current consumption to the marginal utility of wealth. Equation (3) provides the set of pricing restrictions. All returns are now in excess of the riskless interest rate, which is indicated by the use of upper case "R's". $E(\cdot|\Im_t)$ is the expectation of excess returns on security i, $Cov\left(R_{i,t+1}, R_{j,t+1}, |\Im_t\right)$ and $Cov\left(R_{i,t+1}, X_{l,t+1}, |\Im_t\right)$ are, respectively, the covariance between returns on asset i and j, and the covariance between returns on security i and the state variable X_l. Both first and second moments are conditional on the current information set $\Im_t$. $w_{j,t}$ is the optimal wealth share of each risky asset. $\lambda_{M,t} \equiv -J_{WW,t}W_t/J_{W,t}$ is the Arrow-Pratt coefficient of relative risk aversion (provided that $J_{WW,t} < 0$), where $J_{W,t}$ and $J_{WW,t}$ denote the first and second derivatives, respectively, of $J(W, X, t)$ with respect to W. Since $\lambda_{M,t}$ measures how sensitive expected excess returns are to changes in the market risk, it is interpreted as the price of market risk. Similarly, since $Cov\left(R_{i,t+1}, X_{l,t+1}, |\Im_t\right)$ measures the exposure of asset i to the risk stemming from changes in the investment opportunity set, $\lambda_{Fl,t} \equiv -J_{WX_l,t}/J_{W,t}$, $l = 1, \ldots, m$, can be interpreted as the price of intertemporal risk. As before, $J_{WX_l,t}$ is the derivative of the marginal utility of wealth with respect to the state variable X_l. Both $\lambda_{M,t}$ and $\lambda_{Fl,t}$ $\forall l$ are aggregate measures, in that they are harmonic means of the prices of risk of each investor.

Note that the conditional multi-factor model nests the traditional CAPM as a special case: If the marginal utility of wealth is state-independent, i.e. $J_{WX_l,t} = 0$, which happens for agents with one period lives, then Merton's Intertemporal CAPM reduces to the classical CAPM. The second term in equation (3) reflects the need to hedge against adverse shifts in the investment opportunity set. A change in X_l such that future consumption decreases given future wealth represents an "unfavorable" shift in investment opportunities.

As long as $J_{W,t} > 0$ and investors are risk-averse, i.e. $J_{WW,t} < 0$, the price of market risk must always be positive. However, the model does not impose any sign restriction on the prices of intertemporal risk. In particular, if $J_{WX_l,t} > 0$ (< 0), then $\lambda_{Fl,t}$ is negative (positive). When $\lambda_{Fl,t}$ and $Cov\left(R_{i,t+1}, X_{l,t+1}, |\Im_t\right)$ have the same sign, the risk premium required to hold asset i increases; conversely, if the price of intertemporal risk and the covariance between the return on security i and the state variable X_l have different signs, the required total risk premium should decrease. The result is intuitive. Let X_l be a business cycle proxy, with X_l decreasing during contractions and increasing during expansions. When the economy goes into recession, suppose that the marginal utility of wealth decreases, and therefore $\lambda_{Fl,t} < 0$. Assume also that expected returns on asset i are positively correlated with the state variable X_l. In this case, the intertemporal risk premium will turn out to be negative: Investors will require lower compensation, in terms of expected returns, to hold that security. If, instead, asset returns and the state variable show a negative correlation, then the intertemporal risk premium will become positive. This might be the case, at least during some periods, if, like in this research, X_l is the growth rate of industrial production. Since the latter is usually not in phase with, say, equity returns (see, for instance, Hamilton and Lin, 1996, Chauvet and Potter, 1998, and Chauvet, 1999), it can occur that expected asset returns go up during business cycle troughs, in anticipation of a future recovery. This negative correlation may produce a positive intertemporal risk premium.

As with the traditional CAPM, Merton's Intertemporal CAPM suffers from the drawback that it is a partial equilibrium model in its nature. Therefore, it offers little guidance in the choice of pricing factors and is silent about the forces that determine factor risk prices. The identification of the priced state variables is worth a separate study (see, for instance, Chen, Roll, and Ross, 1986, and Fama, 1998). However, Merton's model itself can give some indication (Cochrane, 1999). In a nutshell, Merton's intertemporal asset pricing theory builds a bridge between consumption smoothing and asset returns: A security should pay low average returns if it performs better than other assets during "bad times" and relatively worse during "good times". "Bad times" are periods characterized by a decline in consumption. Variables such as industrial production, oil prices, inflation, business cycle proxies, term structure or interest rates are related to consumption and, therefore, are good candidates for additional priced factors. To keep the analysis empirically tractable, only the growth rate of industrial production will be considered here.

3 Data Diagnosis

The data diagnosis carried out in this section serves two purposes: On one hand, the variables employed in the empirical investigation are discussed. But, more importantly, it will be shown that the considered time series contain some of the features that GARCH processes and regime switching models can capture, making these econometric techniques well-suited for the analysis. Return series show little or no autocorrelation, whereas their squared values are not serially independent, a clear indication of volatility clustering. Besides, unconditional distributions are non-normal and, in particular, they exhibit thick tails as well as significant skewness. Furthermore, all variables are seen to be sensitive to the sign of past shocks, a characteristic which will be taken into account when modelling the GARCH.

In the spirit of Bollerslev *et al.* (1988), a portfolio composed of three assets will be considered: US stocks, 6-month US Treasury bills, and 10-year US government bonds. This allows for the inclusion of a large variety of liquid investment opportunities. All observations are from the last trading day of the month, and cover a period from January 29 1960 to December 31 1998, for a total sample size of 468. All returns are in excess of the 1-month risk-free rate as computed by the Center for Research in Security Prices (CRSP) at the University of Chicago. To measure total (excess) stock returns, R_t^{stock}, the S&P 500 composite index is used, which is market-value-weighted and includes dividends[6]. Yields to maturity on 6-month T-bills are again from CRSP. Next, following Bollerslev *et al.* (1988), these yields are used to compute the associated (excess) holding yields, R_t^{bill}. The one-month holding period returns for 10-year government bonds are, instead, provided directly from CRSP, while the corresponding excess returns, R_t^{bond}, are computed by subtracting from the former the risk-free rate.

The market values of corporate equities are from the *Flows of Funds Accounts of the United States* (Federal Reserve), while the maturity distribution of interest-bearing public debt held by private investors is taken from the *Federal Reserve Bulletin* and the *Treasury Bulletin.* Data on the outstanding public debt are provided in par values and, therefore, need to be converted into market values. This is accomplished by multiplying the different maturity categories of the debt, (i.e. within 1 year, 1 to 5 years, 5 to 10 years, and 10 years and over), by the price indices computed by Cox (1985)[7].

The second priced factor used here is the US growth rate of industrial production (IP), computed with a total production index seasonally adjusted (source: Federal Reserve). The series has the shortcoming that it is lagged by about half a month with respect to the other series employed in the analysis. This, however, should not constitute a problem due to the lower volatility of macro-variables when compared to financial data.

Descriptive statistics for the three assets and the industrial production growth rate are given in Table 1. Panel 1A shows that all distributions are skewed and leptokurtik at the 1% significance level, a clear indication of non-normality. This is confirmed by the Jarque-Bera normality test. Stock excess returns and excess holding yields on 10-year government bonds exhibit no autocorrelation, as indicated by the Ljung-Box statistics. As for the excess holding yields on 6-month T-bills and the growth rate of industrial production the test statistics indicate autocorrelation at the 1% significance

[6]I thank S&P - DRI in Milan, and particularly Ilaria Brioschi, who kindly provided these data.

[7]The index series published in Cox (1985) ends in December 1984. I thank William Cox from the Federal Reserve Bank of Dallas who provided me with the updated indices.

level. Finally, augmented Dickey-Fuller unit root tests on the three assets plus the extra factor reject the null of non-stationarity at the 1% significance level.

In Panel 1B autocorrelation functions are reported for lags from 1 to 12. The three assets show little serial autocorrelation, since only ρ_5 for stocks and ρ_1 for T-bills are significantly different from zero. As for IP, the autocorrelation functions for lags from 1 to 12 confirm the results seen with the Ljung-Box test statistics. The series is serially autocorrelated up to the order four, a feature that will be taken into account when modeling its dynamic.

Unconditional correlations among assets (Panel 1C) are always significant at the 1% level. As expected, there is a positive correlation between stocks and bonds, on one hand, and bills and bonds, on the other. Interestingly, the correlation between stocks and bills is negative. The magnitude of the asset correlation coefficients (never above 0.3) suggest that diversification across asset classes is beneficial. Note also that all securities are inversely correlated with the growth rate of industrial production, though the correlation coefficient between equities and IP is not significant. As for equities this is due to the fact that the series, though linked, are not in phase (see, among others, Hamilton and Lin, 1996, Chauvet and Potter, 1998, and Chauvet, 1999). On the other hand, the negative relationship between T-bill yields and IP finds its justification in standard Keynesian models.

Cross correlations for 6 lags and leads between each pair of assets as well as for each asset and the growth rate of industrial production generate a total of 72 figures. Here, only cross correlations between excess stock market returns and the other two assets plus IP are reported, not only for 6 lags and leads but also contemporaneously, for a total of 39 values. Out of 72 non-synchronous cross correlations, only 21 (including the 14 shown in panel 1D) turn out to be significant. This means that leading or lagging spillovers in excess returns are quite rare, and therefore there is no need to model them in the mean equation specifications.

While excess asset returns as well as the growth rate of industrial production in levels show little autocorrelation, excess *squared* asset returns and IP exhibit a second order dependence that a GARCH process and regime switching models should be able to capture. As shown in Panel 1E, serial correlation is particularly pronounced for T-bills and bonds.

Out of the 6 unconditional contemporaneous correlations between squared time series, 3 are significant. Panel 1F indicates that volatility spillovers

Table 1
Descriptive statistics for excess returns on the stock market, holding yields on 6-month T-bills and 10-year government bonds, and the growth rate of industrial production

Panel 1A: Distributional statistics

	R_t^{stock}	R_t^{bill}	R_t^{bond}	IP
Mean	0.566	0.479	0.169	0.276
Min.	-21.891	-6.272	-7.615	-4.158
Max.	15.994	15.513	9.376	3.402
Std. Dev.	4.251	1.664	2.248	0.818
Skew.	-0.392^{**}	2.449^{**}	0.334^{**}	-0.532^{**}
Kurt.	5.142^{**}	21.856^{**}	4.444^{**}	5.992^{**}
J-B	101.494^{**}	7400.849^{**}	49.350^{**}	196.656^{**}
L-B$_{12}$	8.823	73.907^{**}	16.163	129.870^{**}
D-F	-8.766^{**}	-9.341^{**}	-9.756^{**}	-7.690^{**}

** denotes 1% significance level.
R_t^{stock}, R_t^{bill}, R_t^{bond}, and IP represent the excess returns on the stock market, the holding yields on 6-month T-bills and 10-year government bonds, and the growth rate of industrial production.
Mean, min., max. and standard deviation are in %.
The significance level for skewness (skew.) and excess kurtosis (kurt.) is based on test statistics developed by D'Agostino, Belanger and D'Agostino (1990). The Jarque-Bera (J-B) test for normality combines excess skewness and kurtosis, and is asymptotically distributed as χ_m^2 with $m = 2$ degrees of freedom. The Ljung-Box$_m$ (L-B$_m$) statistics tests the null hypothesis that all autocorrelation coefficients are simultaneously equal to zero up to m lags; it is asymptotically distributed as χ_m^2. It has been chosen $m = 12$ for which the critical values at 95% and 99% confidence level are 21.026 and 26.217, respectively.
D-F is the augmented Dickey-Fuller unit root test statistics; 5% and 1% MacKinnon critical values are -3.421 and -3.982, respectively.

Panel 1B: Autocorrelations of excess returns and the growth rate of industrial production

	R_t^{stock}	R_t^{bill}	R_t^{bond}	IP
ρ_1	0.024	0.282**	0.087	0.344**
ρ_2	−0.029	0.041	−0.001	0.256**
ρ_3	0.004	−0.009	−0.067	0.215**
ρ_4	−0.002	0.065	0.041	0.166**
ρ_5	0.093*	0.008	−0.009	0.048
ρ_6	−0.072	−0.059	0.064	0.031
ρ_{12}	0.022	−0.085	0.009	−0.049

* and ** denote 5% and 1% significance levels, respectively.
ρ_k is the autocorrelation function at lag k.

Panel 1C: Unconditional correlations of excess returns and the growth rate of industrial production

	R_t^{stock}	R_t^{bill}	R_t^{bond}	IP
R_t^{stock}	1.000	−0.141**	0.297**	−0.044
R_t^{bill}		1.000	0.166**	−0.167**
R_t^{bond}			1.000	−0.160**
IP				1.000

** denote 1% significance level, respectively.

Panel 1D: Cross correlations of excess returns between the stock market index, the other two assets and the growth rate of industrial production

	R_t^{bill}	R_t^{bond}	IP
t_{-6}	0.136^{**}	-0.040	0.131^{**}
t_{-5}	0.053	-0.006	0.148^{**}
t_{-4}	0.040	0.035	0.158^{**}
t_{-3}	0.054	-0.068	0.135^{**}
t_{-2}	0.200^{**}	-0.049	0.049
t_{-1}	0.183^{**}	-0.108^{*}	0.011
t_0	-0.141^{**}	0.297^{**}	-0.044
t_1	-0.154^{**}	0.145^{**}	-0.062
t_2	-0.038	0.093^{*}	-0.095^{*}
t_3	0.016	-0.001	-0.045
t_4	-0.056	0.034	-0.051
t_5	-0.083	0.114^{**}	-0.078
t_6	-0.079	0.102^{*}	-0.079

* and ** denote 5% and 1% significance levels, respectively.
t_{-m} represents the cross correlation at lag m, while t_m is the cross correlation at lead m.

Panel 1E: Autocorrelations of squared excess returns and growth rate of industrial production

	R_t^{stock}	R_t^{bill}	R_t^{bond}	IP
ρ_1	0.098^{*}	0.304^{**}	0.173^{**}	0.317^{**}
ρ_2	0.082	0.101^{*}	0.286^{**}	0.122^{**}
ρ_3	0.097^{*}	0.068	0.193^{**}	0.068
ρ_4	0.045	0.092^{*}	0.186^{**}	0.039
ρ_5	0.019	0.063	0.082	-0.021
ρ_6	0.049	0.147^{**}	0.135^{**}	0.041
ρ_{12}	0.022	0.120^{**}	0.184^{**}	0.061

* and ** denote 5% and 1% significance levels, respectively.
ρ_k is the autocorrelation function at lag k.

Panel 1F: Unconditional correlations of squared excess returns and growth rate of industrial production

	R_t^{stock}	R_t^{bill}	R_t^{bond}	IP
R_t^{stock}	1.000	0.122**	0.239**	0.062
R_t^{bill}		1.000	0.179**	0.017
R_t^{bond}			1.000	−0.023
IP				1.000

** denotes 1% significance level.

between different financial markets are important. Therefore, a diagonal GARCH parametrization is not fully appropriate for this sample. It will nevertheless be used to economize parameters to estimate and facilitate convergence.

Only 15 out of 72 non-contemporaneous cross correlations of the squared time series up to 6 lags and leads turn out to be significant. Thus, volatility spillovers at different times do not need to be modelled. Note that panel 1G shows only cross correlations between squared stock market returns and the remaining variables.

In addition to descriptive statistics, the diagnostic test statistics proposed by Engle and Ng (1993) are carried out. These tests examine whether the raw data volatility responds asymmetrically to the sign of past unconditional shocks, without imposing any *a priori* volatility model. Since the growth rate of industrial production also shows significant skewness, it is natural to apply the Engle and Ng (1993) test statistics to that time series as well. Define $v_{i,t} = R_{i,t} - \mu_i$, $i = stock, bill, bond$, where $R_{i,t}$ is the (excess) return on asset i, and μ_i its unconditional mean. When the test is conducted on IP, one has $v_{IP,t} = IP_t - \mu_{IP}$. The square of $v_{i,t}$ can be considered a rough measure of the unconditional volatility. Thereafter, run the following regression:

$$v_{i,t}^2 = \alpha_{i0}+\alpha_{i1}I\left(v_{i,t-1}<0\right)+\alpha_{i2}I\left(v_{i,t-1}<0\right)v_{i,t-1}+\alpha_{i3}I\left(v_{i,t-1}>0\right)v_{i,t-1}+e_{i,t},\ \forall i, \tag{4}$$

where α_{i0}, α_{i1}, α_{i2}, and α_{i3}, are constant coefficients; $I(\cdot)$ denotes the indicator function which is equal to one if the argument is true and zero otherwise; and, finally, $e_{i,t}$ is the residual. The t-ratios for α_{i1}, α_{i2}, and α_{i3} are the sign bias, the negative size bias, and the positive size bias test statistics, respectively. The sign bias test detects whether positive and negative (return)

Panel 1G: Cross correlations of squared excess returns between the stock market index, the other two assets and the growth rate of industrial production

	R_t^{bill}	R_t^{bond}	IP
t_{-6}	0.001	0.053	0.041
t_{-5}	0.008	−0.028	0.086
t_{-4}	0.068	−0.027	0.080
t_{-3}	0.017	0.075	0.176**
t_{-2}	0.097	0.048	0.193**
t_{-1}	0.070	0.036	0.084
t_0	0.122**	0.239**	0.062
t_1	0.029	0.076	0.019
t_2	0.024	0.041	−0.002
t_3	−0.001	0.037	−0.024
t_4	0.012	0.029	−0.044
t_5	0.027	0.008	−0.030
t_6	−0.018	0.052	0.021

** denotes 1% significance level.

t_{-m} represents the cross correlation at lag m, while t_m is the cross correlation at lead m.

Table 2

Diagnostic tests for volatility asymmetry of excess returns on the stock market, holding yields on 6-month T-bills, 10-year government bonds, and the growth rate of industrial production

The estimated model is:

$$v_{i,t}^2 = \alpha_{i0}+\alpha_{i1}I\left(v_{i,t-1}<0\right)+\alpha_{i2}I\left(v_{i,t-1}<0\right)v_{i,t-1}+\alpha_{i3}I\left(v_{i,t-1}>0\right)v_{i,t-1}+e_{i,t},\ \forall i.$$

	α_{i0}	α_{i1}	α_{i2}	α_{i3}	LM	p-value
$R_{stock,t}$	11.338 (3.592)	7.747 (5.111)	-1.873 (0.802)	0.003 (0.907)	22.428	0.000
$R_{bill,t}$	-0.636 (1.026)	1.990 (1.424)	-1.486 (0.781)	3.156 (0.492)	41.296	0.000
$R_{bond,t}$	3.367 (0.984)	-0.487 (1.299)	-0.984 (0.430)	1.360 (0.385)	20.629	0.000
IP	0.220 (0.110)	-0.405 (0.171)	-1.684 (0.143)	0.403 (0.135)	122.009	0.000

R_t^{stock}, R_t^{bill}, R_t^{bond}, and IP represent the excess returns on the stock market, the holding yields on 6-month T-bills and 10-year government bonds, and the industrial production growth rate.
Standard errors are shown in parentheses.
The t-ratios for α_{i1}, α_{i2}, and α_{i3} are the sign bias, the negative size bias, and the positive size bias test statistics, respectively (Engle and Ng, 1993).
The Lagrange Multiplier statistic (LM) tests the null $H_0 : \alpha_{i1} = \alpha_{i2} = \alpha_{i3} = 0$. It is asymptotically distributed as a χ_m^2 with $m = 3$ degrees of freedom for which the critical values at 95% and 99% confidence level are 7.815 and 11.345, respectively.

shocks have an effect on $v_{i,t}^2$. The negative size bias test investigates the effect of large and small *negative* (return) shocks on $v_{i,t}^2$. Finally, the positive size bias test focuses on the impacts that large and small *positive* (return) shocks may have on $v_{i,t}^2$. A Lagrange Multiplier (LM) joint test for the null $H_0 : \alpha_{i1} = \alpha_{i2} = \alpha_{i3} = 0$ can also be computed. Such test statistics is equal to T (the sample size) times the R-squared from the regression. The LM test statistic has the standard limiting distribution and, in particular, is distributed as a chi-square with three degrees of freedom. Table 2 reports parameter estimates from equation (4) as well as the LM test statistic for each asset and the additional pricing factor.

Not surprisingly, when analyzing stocks, α_2 is the only significant parameter at 5% level, suggesting that the volatility of equity returns is sensitive to the negative rather than the positive sign of past shocks. The opposite occurs for T-bills: α_3 is the only significant parameter in that regression, indicating the relative importance of positive rather than negative past news. As for government bonds, both α_2 and α_3 are significant at 5% level, suggesting that negative and positive innovations have a different impact on the volatility of that asset.

The fact that positive past innovations have, in relative terms, a negligible impact on equity volatility, being the probability of rejecting the null $H_0 : \alpha_3 = 0$ equal to 0.31%, can be explained by the feedback volatility hypothesis together with the leverage effect. The combination of the two mechanisms attributes more importance to negative than to positive shocks of the same magnitude. Even though the feedback volatility hypothesis probably does not provide a satisfactory explanation for the asymmetric volatility phenomenon of T-bills and government bonds, it offers, nevertheless, a theoretical underpinning. As for T-bills, the probability of rejecting the null $H_0 : \alpha_2 = 0$ is equal to 94.24%. Therefore, even though α_2 turns out to be consistent with zero at statistical conventional levels, the volatility feedback hypothesis plays its role, since T-bill current volatility exhibits a certain sensitivity to negative past shocks. Why positive innovations are more important than negative remains to be investigated and more satisfactory theoretical foundations are called for. As for government bonds, since the null hypothesis $H_0 : \alpha_2 = 0$ is rejected with a probability equal to 97.74%, negative past innovations increase current volatility. In addition to that, since no leverage effect can be advocated for this security, anticipated positive shocks raise volatility as well, but with a different weight.

The growth rate of industrial production shows a behavior similar to bonds, in that both α_2 and α_3 are significant at 5% level. The asymmetric volatility effect for this series does not find an appropriate justification in the theories put forth by Black (1976) and Christie (1982) and Campbell and Hentschell (1992) and, for the time being, can be considered a mere statistical phenomenon.

The LM test statistics rejects the null hypothesis $H_0 : \alpha_{i1} = \alpha_{i2} = \alpha_{i3} = 0$ for the three securities as well as the industrial production growth rate at 1% significance level, indicating that volatility responds asymmetrically to past news. These results are consistent with the fact that asset returns as well as the state variable are significantly skewed: The variance of a distribution which is not centered is different from the variance of a symmetrical distribution, and, in particular, can respond differently to the sign of past innovations. Harvey and Siddique (1999) estimate a conditional mean equation for stock returns including an autoregressive conditional skewness function through a noncentral t-distribution. Interestingly, the inclusion of skewness substantially reduces the asymmetry in conditional variance, indicating the existence of a link between skewness and asymmetric volatility. More importantly, the Engle and Ng (1993) test statistics point out that the asymmetric effect on volatility is not the same for each asset, a feature that has to be taken into account when modeling the time evolution of conditional second moments.

4 Empirical Methods

Under the assumption that investors are rational, the set of pricing restrictions (3) provides the following statistical model:

$$\mathbf{R}_{t+1} = \lambda_{M,t} \sum_{j=1}^{n} \mathbf{h}_{j,t+1} w_{j,t} + \sum_{l=1}^{m} \lambda_{Fl,t} \mathbf{h}_{n+l,t+1} + \boldsymbol{\varepsilon}_{M,t+1}, \tag{5}$$

where $\mathbf{R}_{t+1}$ represents the $n \times 1$ vector of security excess returns, $\mathbf{h}_{j,t+1}$ the $n \times 1$ vector of conditional variance/covariances of each asset with itself/others, $\mathbf{h}_{n+l,t+1}$ the $n \times 1$ vector of conditional covariances of each security with the state variables, and $\boldsymbol{\varepsilon}_{M,t+1}$ the $n \times 1$ vector of conditional error terms. As before, $\lambda_{M,t}$ and $\lambda_{Fl,t}$ are, respectively, the prices of market and intertemporal risk, while $w_{j,t}$ is the optimal wealth share of each risky asset.

The theoretical model does not impose any restriction on the parametrization of the dynamic of the additional factors. Therefore one can choose a functional form of the kind:

$$\mathbf{F}_{t+1} = \mathbf{K}\mathbf{y}_t + \boldsymbol{\varepsilon}_{F,t+1}, \tag{6}$$

where $\mathbf{F}_{t+1}$ is the $m \times 1$ vector of priced factors, $\mathbf{y}_t$ is a $k \times 1$ vector of predetermined variables which have predictive power with respect to factors, $\mathbf{K}$ is the associated $m \times k$ matrix of parameters, and $\boldsymbol{\varepsilon}_{F,t+1}$ is the $m \times 1$ vector of conditional error terms.

To accommodate the non-normality typical of financial time series, and in particular the leptokurtosis, the $(n + m) \times 1$ disturbance vector, $\boldsymbol{\varepsilon}_{t+1} = [\boldsymbol{\varepsilon}_{M,t+1}\ \boldsymbol{\varepsilon}_{F,t+1}]'$, is assumed to be conditionally Student t-distributed (Bollerslev, 1987)

$$\boldsymbol{\varepsilon}_{t+1} \left| \Im_t \sim Student - t \ \left(0, \mathbf{H}_{t+1}, \nu\right) \right.,$$

where ν is the parameter indicating the degrees of freedom. $\mathbf{H}_{t+1}$ is the $(n + m) \times (n + m)$ conditional covariance matrix of asset returns and priced factors. Note that $\mathbf{h}_{j,t+1}$, $j = 1, \ldots, n$, and $\mathbf{h}_{n+l,t+1}$, $l = 1, \ldots, m$, are, respectively, the first n and the last m columns of $\mathbf{H}_{t+1}$. Economic theory does not suggest any hypothesis about conditional second moment evolution, nor about their relationship with economic fundamentals. Therefore, one has to rely on *ad hoc* assumptions and on specific statistical models. GARCH processes are among the most widely used parametrizations to model conditional second moments. It is assumed that the conditional covariance matrix follows a multivariate GARCH(1,1) process, according to Engle and Kroner (1995):

$$\mathbf{H}_{t+1} = \mathbf{V}'\mathbf{V} + \mathbf{A}'\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{A} + \mathbf{B}'\mathbf{H}_t\mathbf{B}, \tag{7}$$

where $\mathbf{V}$, $\mathbf{A}$, and $\mathbf{B}$ are $(n+m) \times (n+m)$ matrices of unknown parameters[8]. If, for simplicity, $\mathbf{A}$ and $\mathbf{B}$ are assumed to be diagonal, the process (7) becomes

$$\mathbf{H}_{t+1} = \mathbf{V}'\mathbf{V} + \mathbf{aa}' \odot \boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' + \mathbf{bb}' \odot \mathbf{H}_t, \tag{8}$$

where $\mathbf{a}$ and $\mathbf{b}$ are $(n+m) \times 1$ vectors of parameters which include the diagonal elements of $\mathbf{A}$ and $\mathbf{B}$, respectively, while $\odot$ denotes the Hadamard (element by element) matrix product. The number of unknown parameters in equation (8) is $(n+m)\left[(n+m)+5\right]/2$, which can easily become prohibitive during the estimation procedure if the number of assets and/or factors is large. Moreover, a simpler GARCH parametrization would facilitate convergence. Following Ding and Engle (1994), an even more parsimonious representation is therefore adopted. Under the assumption of covariance stationarity, the unconditional covariance matrix of the residuals implied by equation (8) will be $\mathbf{H}_0 = \mathbf{V}'\mathbf{V} \odot \left(\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}'\right)^{-1}$. Therefore, system (8) reduces to

$$\mathbf{H}_{t+1} = \mathbf{H}_0 \odot \left(\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}'\right) + \mathbf{aa}' \odot \boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' + \mathbf{bb}' \odot \mathbf{H}_t, \tag{9}$$

where $\mathbf{1}$ represents the unit vector. $\mathbf{H}_0$, though unknown, can be evaluated recursively during the optimization procedure, according to the methodology suggested by De Santis and Gerard (1997, 1998a)[9]. With this parametrization, the number of unknown parameters reduces to $(n+m) \times 2$, greatly facilitating the estimation of non-linear models.

The GARCH process represented in equation (9) is symmetric: It does not accommodate the fact that the volatility of asset returns and the growth rate of industrial production is sensitive to the sign of past innovations. However, as reflected by the Engle and Ng's (1993) tests statistics (see Section 3), volatility responds asymmetrically to negative and/or positive shocks. Therefore, conditional second moments should be modeled in such a way

[8]Note that $\mathbf{V}$ is an upper triangular matrix. Therefore it only contains $\left[(n+m)(n+m+1)\right]/2$ unknown parameters. $\mathbf{V}'\mathbf{V}$ is, in fact, a Cholesky decomposition of the GARCH constant term, which ensures the positive definiteness of the process.

[9]At the first iteration $\mathbf{H}_0$ is set equal to the unconditional covariance matrix. Then it is updated at the end of each iteration.

that takes into account the asymmetric volatility effect. The next Section discusses how to modify the multivariate GARCH process described by equation (9) in order to capture asymmetry in both conditional variances and covariances.

4.1 Modeling the Asymmetric Effects in Multivariate GARCH Processes

As pointed out, among others, by Nelson (1991), a GARCH process of the kind represented by equation (9) suffers from an important limitation. Although it elegantly captures volatility clustering, it does not allow negative and positive past shocks to have a different effect on future conditional second moments. In other words, only the magnitude, not the sign of lagged innovations determines conditional variances and covariances. Therefore a model that captures the asymmetric responses of conditional second moments should be preferable for asset pricing applications. To better see this, consider, for a moment, a portfolio made of equities only and think, for instance, of a large price drop, like the one that happened in October 1987. If a negative return innovation generates more volatility than a positive return innovation of the same magnitude, a symmetric GARCH process underestimates the conditional volatility which occurs after bad news. Similarly, it overestimates the conditional volatility which follows good news. In CAPM-type models, conditional volatility directly affects risk premia investors require to hold risky assets. But the premia forecast by the traditional GARCH differ from those implied by an asymmetric GARCH, with a consequence of probable asset mis-pricing.

The GARCH literature has devised numerous ways to model asymmetry, particularly in univariate contexts. Examples are given by the quadratic GARCH of Engle (1990) and Sentana (1991), the EGARCH model introduced by Nelson (1991), the qualitative threshold ARCH of Gouriéroux and Monfort (1992), the asymmetric power ARCH of Ding, Granger, and Engle (1993), the Glosten-Jagannathan-Runkle (1993) (GJR) GARCH, the nonlinear asymmetric GARCH of Engle and Ng (1993), the threshold GARCH of Zakoian (1994) thereafter extended by Rabemananjara and Zakoian (1993)[10], the volatility-switching ARCH proposed by Fornari and Mele (1997), the smooth transition GARCH of Lubrano (1998), and the EGARCH of Wu and Xiao (2000), where the impact of return shocks on conditional volatility is estimated nonparametrically. Probably, the most popular asymmetric

[10]The first version of Zakoian's threshold GARCH appeared in a working paper (CREST, INSEE) dated 1990. This explains why the extension of the model due to Rabemananjara and Zakoian (1993) has an earlier date than the original model published only in 1994.

univariate GARCH representations are the EGARCH of Nelson and the GJR GARCH of Glosten, Jagannathan, and Runkle. These two GARCH processes are extended here to a multivariate context starting from the parametrization given by (9). However, to begin with, it is convenient to first briefly discuss the univariate specifications for both the EGARCH and the GJR models.

4.1.1 The Generalized Nelson Model

Let $u_{t+1} = \sqrt{h_{t+1}}\upsilon_{t+1}$ be a model's prediction error, where υ_{t+1} is $i.i.d.$ with zero mean and unit variance. Nelson's (1991) univariate EGARCH process for the evolution of the conditional variance of u_{t+1}, h_{t+1}, is as follows:

$$\ln h_{t+1} = \alpha_0 + \sum_{i=1}^{\infty} \alpha_i \left\{ \gamma \left[|\upsilon_{t+1-i}| - E\left|\upsilon_{t+1-i}\right| + \zeta \upsilon_{t+1-i} \right] \right\} + \sum_{j=1}^{q} \beta_j \ln h_{t+1-j}, \;\; \alpha_1 \equiv 1. \tag{10}$$

Nelson's specification has at least two advantages over a standard GARCH model. First, since equation (10) describes the log of h_{t+1}, the variance itself is always positive, without imposing any positivity constraint on the parameters. Second, the EGARCH parametrization accommodates the asymmetric relation between asset returns and volatility, since the conditional variance depends on both the size *and* the sign of υ_{t+1-i}. To better see that, assume that the GARCH process (10) is of order $(1, 1)$. Similarly to the traditional GARCH representation, the term $\gamma\left[|\upsilon_t| - E\left|\upsilon_t\right|\right]$ captures the magnitude effect. This becomes clear if one assumes that $\gamma > 0$ and $\zeta = 0$. In this case the $\gamma\left[|\upsilon_t| - E\left|\upsilon_t\right|\right]$ term raises $\ln h_{t+1}$ when the magnitude of market movements is large whereas lowers it when it is small. Moreover, when $\zeta = 0$, the conditional volatility responds symmetrically to the lagged return innovations. In particular, a positive surprise (which occurs when the magnitude of υ_{t+1} is larger than zero) has the same effect on volatility than a negative surprise of the same magnitude (which happens if the reverse occurs, i.e. υ_{t+1} is smaller than zero). The term $\zeta\upsilon_t$ allows the conditional variance to respond asymmetrically to the sign of lagged return innovations. To see that, suppose that $\gamma = 0$ and $\zeta < 0$. In this case a positive return innovation actually decreases volatility, while a negative return innovation increases it.

The GARCH model described in (10) can be extended to a multivariate context, when, for instance, a system of mean equations has to be estimated and the conditional covariance matrix of the error term vector is supposed to follow an asymmetric GARCH process. In the literature there exist already examples of EGARCH applied to systems of equations: See, among the others, Braun, Nelson, and Sunier (1995), Koutmos and Booth (1995), Koutmos (1996), Booth, Martikainen, and Tse (1997), and Scruggs (1998).

In all these studies the conditional variances follow equations of the type described in (10), but the conditional covariances do not explicitly account for any asymmetric effect. Apart from the research of Braun *et al.*, conditional covariances, $h_{ij,t+1}$, are modeled in the spirit of Bollerslev (1990). That is, $h_{ij,t+1} = \rho_{ij}\sqrt{h_{i,t+1}}\sqrt{h_{j,t+1}}$, $\forall i$ and j, where ρ_{ij} is the conditional correlation coefficient between asset i and j and it is assumed to be constant over the sample period, while $\sqrt{h_{i,t+1}}$ and $\sqrt{h_{j,t+1}}$ are, respectively, the conditional standard deviations of security i and j. Braun *et al.*, who analyze a portfolio of two assets, model the second moment matrix by splitting it into three pieces: The two conditional variances associated with each security and the conditional beta. Also in this case conditional covariances do not exhibit explicit asymmetric effects. However, as pointed out by Kroner and Ng (1998), if, due to the asymmetric volatility effect, expected returns on one asset change, the covariance between returns on that asset and returns on assets which have possibly not experienced such an effect should also change. Therefore, the following multivariate GARCH specification is proposed, with the scope of explicitly capturing asymmetric comovements of asset returns:

$$\mathbf{H}_{t+1} = \mathbf{H}_0 \odot (\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}') + \mathbf{aa}' \odot \eta(\boldsymbol{\upsilon}_t)\,\eta(\boldsymbol{\upsilon}_t)' + \boldsymbol{\zeta\zeta}' \odot \boldsymbol{\upsilon}_t\boldsymbol{\upsilon}_t' + \mathbf{bb}' \odot \mathbf{H}_t, \tag{11}$$

The process described in (11), which one can name "generalized Nelson model", is built on the symmetric GARCH parametrization represented in (9). Like the term $\gamma\left[|\upsilon_t| - E|\upsilon_t|\right]$ in equation (10), the term $\mathbf{aa}' \odot \eta(\boldsymbol{\upsilon}_t)\,\eta(\boldsymbol{\upsilon}_t)'$ accommodates the size effect. In particular $\eta(\boldsymbol{\upsilon}_t) = |\boldsymbol{\upsilon}_t| - E|\boldsymbol{\upsilon}_t|$, where $\boldsymbol{\upsilon}_t$ is an $(n+m) \times 1$ vector, whose generic element is $\upsilon_{i,t} = \varepsilon_{i,t}/\sqrt{h_{i,t}}$ $\forall i$, i.e. the ith conditional standardized innovation. The term $\boldsymbol{\zeta\zeta}' \odot \boldsymbol{\upsilon}_t\boldsymbol{\upsilon}_t'$, which in the univariate representation simplifies to $\zeta\upsilon_t$, accommodates the sign effect for both conditional variances *and* covariances. The matrix $\boldsymbol{\zeta\zeta}' \equiv \left[\zeta_{ij}\right]$ is defined as follows: $\zeta_{ii} = \zeta_i$ $\forall i$, and $\zeta_{ij} = \zeta_i\zeta_j$ $\forall i \neq j$. Similarly, for the matrix $\boldsymbol{\upsilon}_t\boldsymbol{\upsilon}_t' \equiv [\upsilon_{ij,t}]$ the elements that lie on the main diagonal are defined as $\upsilon_{ii,t} = \upsilon_{i,t}$ $\forall i$, while the elements off the main diagonal are $\upsilon_{ij,t} = \upsilon_{i,t}\upsilon_{j,t}$, $\forall i \neq j$. The (i,j)th element of the $\mathbf{H}_{t+1}$ matrix described in (11) is:

$$h_{ii,t+1} = h_{ii,0}(1 - a_i a_i - b_i b_i) + a_i a_i \eta(\upsilon_{i,t})\,\eta(\upsilon_{i,t}) + \zeta_i\upsilon_{i,t} + b_i b_i h_{ii,t} \ \forall i,$$

$$h_{ij,t+1} = h_{ij,0}(1 - a_i a_j - b_i b_j) + a_i a_j \eta(\upsilon_{i,t})\,\eta(\upsilon_{j,t}) + \zeta_i\zeta_j\upsilon_{i,t}\upsilon_{j,t} + b_i b_j h_{ij,t} \ \forall i \neq j.$$

This parametrization allows to capture the asymmetric effect on conditional variances and covariances. Assume, for instance, that $\zeta_i < 0$ (> 0). If, *ceteris paribus*, $\upsilon_{i,t} < 0$, volatility will increase (decrease), while if $\upsilon_{i,t} > 0$, volatility will decrease (increase). The sign effect on each covariance will be given by $\zeta_i \zeta_j \upsilon_{i,t} \upsilon_{j,t}$, $\forall i \neq j$.

The last issue to discuss concerns the assumption about the distribution taken by $\boldsymbol{\varepsilon}_{t+1}$, since the value assumed by $E\left|\boldsymbol{\upsilon}_t\right|$ will depend on this choice. Nelson (1991) proposed the use of the Generalized Error Distribution (GED). As mentioned above, residuals here are assumed to be conditionally Student t-distributed. Therefore $E\left|\upsilon_{i,t}\right| = (2/\pi)^{1/2}\left[\Gamma\left(\nu-1\right)/2\right]/\Gamma\left(\nu/2\right)\ \forall i$, where, as before, $\Gamma\left(\cdot\right)$ is the gamma function and ν the degrees of freedom parameter[11].

4.1.2 The Generalized GJR Model

The GJR univariate asymmetric GARCH model has a simpler structure than Nelson representation. As before, let $u_{t+1} = \sqrt{h_{t+1}}\upsilon_{t+1}$ be a model's prediction error, where υ_{t+1} is $i.i.d.$ with zero mean and unit variance. The GJR parametrization, originally designed to estimate the conditional variance of equities, is formulated as follows:

$$h_{t+1} = \alpha_0 + \sum_{i=1}^{p}\left[\alpha_i u_{t+1-i}^2 + \zeta_i I\left(u_{t+1-i} < 0\right) u_{t+1-i}^2\right] + \sum_{j=1}^{q}\beta_j h_{t+1-j}, \quad (12)$$

where, as before, $I\left(\cdot\right)$ is the indicator function. The only difference with a traditional GARCH is the inclusion of the term $\zeta_i I\left(u_{t+1-i} < 0\right) u_{t+1-i}^2$, $i = 1, \ldots, p$. This term accommodates the asymmetric volatility effect. Assume that $\zeta_i > 0$ (< 0). When $u_{t+1-i} < 0$, $I\left(\cdot\right) = 1$ and volatility actually goes up (down). Conditional volatility, instead, does not change for positive innovations, since $I\left(\cdot\right) = 0$. In a comparison study for daily Japanese TOPIX data, Engle and Ng (1993) find that, of several variance parametric models with the inclusion of the Nelson's (1991) EGARCH, the GJR is the best at parsimoniously capturing the asymmetric effect.

A multivariate version of the univariate GJR parametrization is due to Kroner and Ng (1998). In the spirit of that research, the process (9) is modified in order to capture asymmetry in second moments and the following representation is suggested:

$$\mathbf{H}_{t+1} = \mathbf{H}_0 \odot \left(\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}'\right) + \mathbf{aa}' \odot \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' + \boldsymbol{\zeta\zeta}' \odot \boldsymbol{\eta}_t \boldsymbol{\eta}_t' + \mathbf{bb}' \odot \mathbf{H}_t, \quad (13)$$

The process described in (13), which one can baptize "generalized GJR model", presents an additional term, $\boldsymbol{\zeta\zeta}' \odot \boldsymbol{\eta}_t \boldsymbol{\eta}_t'$, with respect to the representation given by (9). By construction process (13) nests (9) as a special

[11]Notice that as the degrees of freedom parameter ν goes to infinity, a Student t-distribution approaches a standard normal distribution. In this case $E\left|\upsilon_{i,t}\right| = (2/\pi)^{1/2}$ $\forall i$ (Hamilton, 1994).

case: When all elements of the matrix $\boldsymbol{\zeta\zeta}'$ are set equal to zero, equation (13) reduces to (9). The term $\boldsymbol{\zeta\zeta}' \odot \boldsymbol{\eta}_t\boldsymbol{\eta}_t'$ accommodates the asymmetric responses of variances and covariances to past shocks. $\boldsymbol{\zeta\zeta}' \equiv \left[\zeta_{ij}\right] = \zeta_i\zeta_j\ \forall i$ and j is a $(n+m) \times (n+m)$ matrix of parameters[12]. $\boldsymbol{\eta}_t$ is a $(n+m) \times 1$ vector. In the case analyzed here its dimension reduces to 4×1 and each of the elements it contains is modeled as follows:

$$\eta_{stock,t} = -I\left(\varepsilon_{stock,t} < 0\right)\varepsilon_{stock,t}, \tag{14}$$

$$\eta_{bill,t} = I\left(\varepsilon_{bill,t} > 0\right)\varepsilon_{bill,t}, \tag{15}$$

$$\eta_{bond,t} = \begin{cases} \zeta_{bond1} \cdot \varepsilon_{bond,t} & \text{if } \varepsilon_{bond,t} > 0 \\ -\zeta_{bond2} \cdot \varepsilon_{bond,t} & \text{if } \varepsilon_{bond,t} \leq 0 \end{cases}, \tag{16}$$

$$\eta_{IP,t} = \begin{cases} \zeta_{IP1} \cdot \varepsilon_{IP,t} & \text{if } \varepsilon_{IP,t} > 0 \\ -\zeta_{IP2} \cdot \varepsilon_{IP,t} & \text{if } \varepsilon_{IP,t} \leq 0 \end{cases}. \tag{17}$$

This parametrization finds its rationale in the Engle and Ng (1993) tests presented in Section 3 and in the functional form of the (i,j)th element of the $\mathbf{H}_{t+1}$ matrix described in (13):

$$h_{ij,t+1} = h_{ij,0}\left(1 - a_ia_j - b_ib_j\right) + a_ia_j\varepsilon_{i,t}\varepsilon_{j,t} + \zeta_i\zeta_j\eta_{i,t}\eta_{j,t} + b_ib_jh_{ij,t}\ \forall i \text{ and } j.$$

In particular, when modeling $\eta_{i,t}$, $\forall i$, only the effects of past positive and negative shocks which have an associated parameter significant at maximum 5% level, as computed with the Engle and Ng's (1993) test statistics, are taken into account. Stock volatility seems to increase more after negative than positive past shocks. Therefore, $\zeta_{stock}^2\eta_{stock,t}^2$ will be positive any time $\varepsilon_{stock,t} < 0$ and zero otherwise, independently of the sign attached to ζ_{stock}. As for the T-bill volatility, it seems to react more after positive past innovations: $\zeta_{bill}^2\eta_{bill,t}^2$ will be positive any time $\varepsilon_{bond,t} > 0$ and zero otherwise. Government bond volatility, instead, seems to respond differently to positive and negative past surprises. The two parameters ζ_{bond1} and ζ_{bond2} capture the different responses to positive and negative past innovations. Finally, the volatility of the industrial production growth rate shows the same feature of the variance of government bonds, thus it has the same specification. In order to avoid overparametrization, in the matrix $\boldsymbol{\zeta\zeta}'$, ζ_i, for $i = bond$ and IP, is set equal to one and, therefore, only ζ_1 and ζ_2 are estimated[13].

[12]Like matrices $\mathbf{aa}'$ and $\mathbf{bb}'$, the matrix $\boldsymbol{\zeta\zeta}'$ can be thought of as the result of the product of a column times a row vector $\boldsymbol{\zeta}$. Such a vector includes the diagonal elements of an underlying $\mathbf{Z}$ matrix of dimension $(n+m) \times (n+m)$. As such the $\boldsymbol{\zeta\zeta}'$ matrix contains only $n+m$ unknown parameters.

[13]The parametrization suggested in equations (14) and (15) is similar to that of Glosten, Jagannathan, and Runkle (1993), whereas equations (16) and (17) can be seen as a generalization of the Threshold GARCH proposed by Zakoian (1994).

4.1.3 A Comparison between the Generalized Nelson and GJR Models

The generalized Nelson model has the advantage that its structure does not depend on the particular variable considered in the analysis. The generalized GJR specification, instead, has to take into account how volatility of a given variable responds to past positive or negative shocks, and, as such, relies on test statistics. If, for instance, one imposed $\eta_{stock,t} = I(\varepsilon_{stock,t} > 0)\, \varepsilon_{stock,t}$, equity conditional volatility would increase more to past positive than negative innovations of the same magnitude, a possibility which is probably wrong. Therefore, the strategy suggested to have the best fit of the data is to first carry out the Engle and Ng (1993) tests on each of the analyzed variables and then to model each element of the vector $\boldsymbol{\eta}_t$ accordingly. Moreover, the generalized Nelson model does not impose the direction in which conditional volatility has to go after a shock, whereas the generalized GJR model does. Suppose ζ_i turns out to be negative and, for the sake of simplicity, consider only equities. When the generalized Nelson model is employed, the variance will increase after a negative innovation and decrease after a positive shock. When, instead, the GJR model is used, since the Engle and Ng's (1993) test indicates that volatility rises significantly after bad news but it does not react when good news occurs, the possibility that positive shocks can even decrease the conditional variance will be excluded. However, such a possibility can arise and be significant. In other words, one could question the power of Engle and Ng's (1993) test statistics, an analysis which would go beyond the scope of this paper. The disadvantage of the generalized Nelson model is that it does not accommodate for the fact that volatility can increase to both positive and negative past shocks, though with different intensities. This can be captured with the generalized GJR model, as shown by the parametrization proposed in equations (16) and (17).

Finally, notice that, in general, univariate GARCH representations need some non-negativity requirements on the parameters of the process to ensure a positive conditional variance. As seen before, Nelson's (1991) specification does not require to impose any constraint on the GARCH parameters since the variance itself will be positive regardless of the sign of the coefficients involved. In a multivariate context, it is required that the conditional covariance matrix be positive definite. A sufficient condition to ensure positive definiteness for the symmetric Ding-Engle representation (9) is that the three addenda $\mathbf{H}_0 \odot (\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}')$, $\mathbf{aa}' \odot \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'$, and $\mathbf{bb}' \odot \mathbf{H}_t$ are positive definite. The last two components are positive definite by construction (see Ding and Engle, 1994, for a proof). As for the first one, it is positive definite if both $\mathbf{H}_0$ and $(\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}')$ are positive definite. $\mathbf{H}_0$ is so by definition, since it is set equal to the unconditional covariance matrix at the beginning of the estimation procedure and then updated at the end of each iteration. The non-singularity of $(\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}')$, instead, is not guaranteed in the estimation procedure. Both the generalized Nelson and GJR models, which are built on the process (9), share the same limitation. In practice, though, the conditions under which the matrix $(\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}')$ does not have full rank are quite restrictive, and therefore, one can use the Ding-Engle representation. In addition to that, however, the generalized Nelson model may show another potential problem. The way in which the term $\boldsymbol{\zeta}\boldsymbol{\zeta}' \odot \boldsymbol{\upsilon}_t\boldsymbol{\upsilon}_t'$ is parametrized might be a further source of non-positive definiteness: Its off-diagonal elements can turn out to be too big relative to the diagonal terms, thus causing $\boldsymbol{\zeta}\boldsymbol{\zeta}' \odot \boldsymbol{\upsilon}_t\boldsymbol{\upsilon}_t'$ to be nonpositive definite. In the generalized GJR model, instead, the component aimed at capturing the asymmetric effects, $\boldsymbol{\zeta}\boldsymbol{\zeta}' \odot \boldsymbol{\eta}_t\boldsymbol{\eta}_t'$, is again positive definite by construction. Because of that, the latter specification can turn out to be preferred to the former.

4.2 The Switching Prices of Risk

Merton's (1973) Intertemporal CAPM will first be estimated assuming that the prices of market and intertemporal risk are constant. Next, prices will be allowed to vary over time, but take on only two values. In particular, the coefficient of risk aversion will be equal to either λ_{M1} or λ_{M2}, while the prices of intertemporal risk λ_{Fl1} or λ_{Fl2}, $\forall l$. Let $S_t^M \in \{1,2\}$ and $S_t^F \in \{1,2\}$ represent the state of investors' preferences at time t, with respect to the price of market and intertemporal risk, respectively. At each point in time the price of market risk is

$$\lambda_{Mt} = \begin{cases} \lambda_{M1}, & \text{if } S_t^M = 1, \\ \lambda_{M2}, & \text{if } S_t^M = 2, \end{cases} , \tag{18}$$

while the prices of intertemporal risk are

$$\lambda_{Flt} = \begin{cases} \lambda_{Fl1}, & \text{if } S_t^{Fl} = 1, \\ \lambda_{Fl2}, & \text{if } S_t^{Fl} = 2, \end{cases} \quad \forall l. \tag{19}$$

The time evolution of the unobserved state variable S_t^u, $u = M, Fl$, $\forall l$, is assumed to follow a first-order Markov chain:

$$\left.\begin{array}{l} P(S_t^u = 1 \left| S_{t-1}^u = 1 \right.) = p^u \\ P(S_t^u = 2 \left| S_{t-1}^u = 1 \right.) = 1 - p^u \\ P(S_t^u = 2 \left| S_{t-1}^u = 2 \right.) = q^u \\ P(S_t^u = 1 \left| S_{t-1}^u = 2 \right.) = 1 - q^u \end{array}\right. , \quad u = M, Fl, \forall l. \tag{20}$$

where p^u and q^u represent the conditional probabilities of remaining in the past state. Therefore, $1 - p^u$ and $1 - q^u$ are the conditional probabilities of switching between states.

To simplify the estimation, only one priced factor is chosen, in addition to the market risk. Therefore, together with the coefficient of risk aversion, there is only one price of intertemporal risk to consider. As a consequence, there will be only two unobserved state variables, S_t^M and S_t^F, which will govern the time evolution of λ_{Mt} and λ_{Ft}, $\forall t$, respectively. In the spirit of Hamilton and Lin (1996) and Susmel (1998), three possible cases arise from the combination of S_t^M and S_t^F. Case i): Independent states, i.e. shifts in the price of market risk are completely unrelated to the factors driving the price of intertemporal risk, therefore S_t^M is independent of S_τ^F $\forall t$ and τ. Case ii): Common states, i.e. shifts in the price of market and intertemporal risk are determined by the same factors, thus $S_t^M = S_t^F = S_t$ $\forall t$. Case iii): Related States, i.e. forces that govern the market and intertemporal risk are the same, but are not in phase. Two sub-cases need to be considered: iiia) The coefficient of risk aversion changes before the intertemporal price of risk, i.e. the price of market risk leads the other price, thus $S_t^F = S_{t-1}^M$ $\forall t$. And iiib) The causality is reversed, namely $S_t^M = S_{t-1}^F$ $\forall t$[14].

Cappiello (1999) estimates a two-factor trivariate Intertemporal CAPM using the same assets employed here, but a different state variable to hedge against adverse changes in the investment opportunity set[15]. That research finds that the model with common states outperforms the others in terms of forecasting. On the grounds of those results, even if here the second priced factor is different, only the common state case is considered. Indeed, since λ_M and λ_{Fl} depend on first and second derivatives of the same derived utility function (see Section 2), it is plausible that their evolution over time is governed by the same latent variable.

When case ii) is analyzed, the transition probabilities are given by equation (20), where, since $u = M = F$, the superscript can be dropped out. The associated switching probability matrix and mean equation are, respectively,

$$\mathbf{P} = \begin{bmatrix} p & 1-q \\ 1-p & q \end{bmatrix}, \tag{21}$$

and

$$\mathbf{R}_{t+1} = \lambda_{M,S_t} \sum_{j=1}^{n} \mathbf{h}_{j,t+1} w_{j,t} + \lambda_{F,S_t} \mathbf{h}_{n+1,t+1} + \boldsymbol{\varepsilon}_{M,t+1}. \tag{22}$$

In Hamilton's (1988, 1989, 1990, 1994) original formulation of regime switching models the transition probabilities are constant over time. This

[14]In previous research myself and Tom Fearnley have developed a program to estimate a conditional static (only the market exposure is considered) single-asset CAPM in which the price of market risk is modelled according to Hamilton's (1988, 1989, 1990, 1994) filter. Given the simple nature of that problem, only one latent variable is analyzed and no need for several model specifications is there required.

[15]The state variable employed by Cappiello (1999) is the default premium, defined as the return on a portfolio of low-grade corporate bonds less the return on a portfolio of long-term government bonds.

assumption is first retained and then relaxed later on. Recent studies suggest that keeping the switching probabilities constant may be an oversimplification and make them dependent upon some conditioning information variables[16]:

$$\begin{array}{l} p_t = \Phi\left(\mathbf{Z}_{t-1}\right) \\ q_t = \Phi\left(\mathbf{Z}_{t-1}\right) \end{array}, \tag{23}$$

where $\mathbf{Z}_{t-1}$ is a vector of predetermined variables which can affect the state transition probabilities. $\Phi(\cdot)$ can be any function whose values lie between zero and one. The functions typically used in the literature are the logistic or the cumulative normal density function. Here it is considered the latter because it makes the probabilities monotonic in the instruments, thus facilitating the interpretation of the parameters. As shown in Cappiello (1999), investors exhibit a higher (lower) degree of risk aversion in correspondence of financial turmoils (quiet periods). Therefore, it is preferable to choose predetermined variables which can contain information on the state of financial markets. Variables such as lagged interest rates, term premium, or dividend yields can be good candidates. A raise in the interest rate, for instance, usually has a negative impact on stock markets. Thus it should increase the likelihood of a switch to the high-risk aversion regime.

Thanks to the transition probabilities it is possible to determine the *ex ante* probabilities, conditional on information available up to time t:

$$P\left(S_{t+1} = k\left|\Im_t; \boldsymbol{\theta}\right.\right) = \sum_{j=1}^{2} P\left(S_{t+1} = k | S_t = j\right) P\left(S_t = j\left|\Im_t; \boldsymbol{\theta}\right.\right), \ k = 1, 2, \tag{24}$$

where $\boldsymbol{\theta}$ is the vector of unknown parameters and $P\left(S_t = j\left|\Im_t; \boldsymbol{\theta}\right.\right)$ represents the filter probability, which indicates the state of the economy at time t. By Bayes' rule, the latter can be written as

$$P\left(S_t = k\left|\Im_t; \boldsymbol{\theta}\right.\right) = \frac{f\left(\mathbf{G}_t\left|S_t = k, \Im_{t-1}; \boldsymbol{\theta}\right.\right) P\left(S_t = k\left|\Im_{t-1}; \boldsymbol{\theta}\right.\right)}{\sum_{j=1}^{2} f\left(\mathbf{G}_t\left|S_t = j, \Im_{t-1}; \boldsymbol{\theta}\right.\right) P\left(S_t = j\left|\Im_{t-1}; \boldsymbol{\theta}\right.\right)}, \ k = 1, 2. \tag{25}$$

where $f\left(\mathbf{G}_t\left|S_t = k, \Im_{t-1}; \boldsymbol{\theta}\right.\right)$ is the likelihood function for the data conditional on the information set and the latent state variable, and $\mathbf{G}_t = \left[\mathbf{R}_t\ \mathbf{F}_t\right]'$ is the vector of asset returns and factors. While ex ante probabilities are commonly used for forecasting, filter probabilities give inference about the

[16]See, for instance, Ghysels (1993), Diebold, Lee and Weinbach (1994), Filardo (1994), Durland and McCurdy (1994), Bekaert and Harvey (1995), Gray (1996), Dahlquist and Gray (1998), and Perez-Quiros and Timmerman (1999).

regime in which the economy is. Given the conditional density function and initial starting values for the *ex ante* probabilities, equations (24) and (25) can be iterated recursively to compute the state probabilities and the parameters of the likelihood function.

The conditional density function is assumed to follow a multivariate normalized Student t-distribution within each state (see Johnson and Kotz, 1972, or Fang, Kotz, and Ng, 1990):

$$f\left(\mathbf{G}_t\left|S_t = k, \Im_{t-1}; \boldsymbol{\theta}\right.\right) =$$

$$= \frac{\Gamma\left(\frac{\nu+g}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left[\pi\left(\nu-2\right)\right]^{\frac{g}{2}}\left|\mathbf{H}_t\left(\boldsymbol{\theta}\right)\right|^{\frac{1}{2}}}\left[1+\frac{\boldsymbol{\varepsilon}_t\left(\boldsymbol{\theta}, S_t = k\right)'\mathbf{H}_t\left(\boldsymbol{\theta}\right)^{-1}\boldsymbol{\varepsilon}_t\left(\boldsymbol{\theta}, S_t = k\right)}{\nu-2}\right]^{-\frac{\nu+g}{2}}, \tag{26}$$

where $g = n + m$ and $\Gamma(\cdot)$ is the gamma distribution. The parameters of the model, including ν, are estimated by maximizing the following likelihood function with respect to $\boldsymbol{\theta}$:

$$L\left(\mathbf{G}_t\left|\Im_{t-1}; \boldsymbol{\theta}\right.\right) = \sum_{t=1}^{T}\ln\phi\left(\mathbf{G}_t\left|\Im_{t-1}; \boldsymbol{\theta}\right.\right), \tag{27}$$

where the density function $\phi(\cdot)$ is a weighted average of the two state-dependent densities:

$$\phi\left(\mathbf{G}_t\left|\Im_{t-1}; \boldsymbol{\theta}\right.\right) = \sum_{j=1}^{2} f\left(\mathbf{G}_t|S_t = j, \Im_{t-1}; \boldsymbol{\theta}\right) P\left(S_t = j\left|\Im_{t-1}; \boldsymbol{\theta}\right.\right). \tag{28}$$

The weights are determined by the *ex ante* probabilities (equation (24)).

The last issue to discuss concerns the presence of switching parameters in the asset mean equations, i.e. λ_{M,S_t} and λ_{F,S_t}. Since the model is estimated with a GARCH-in-Mean technique, switching prices of risk cause the GARCH process to be a function of the entire history of the state variables. Specifically, as for the generalized Nelson model, at time t the conditional (co)variances depend on themselves lagged once and, directly as well as through the $\eta(\cdot)$ function, on conditional standardized innovations. These, in turn, are function of lagged error terms, which depend on the lagged state variables. Therefore the $\mathbf{H}_t$ matrix can be described by the following function:

$$\mathbf{H}_t^{Nelson} = \varphi\left[\boldsymbol{\upsilon}_{t-1}\left(S_{t-2}\right), \mathbf{H}_{t-1}^{Nelson}\right]. \tag{29}$$

Similarly, the conditional GJR covariance matrix depends on itself lagged once and, directly as well as through $\boldsymbol{\eta}_t$, on lagged error terms, which are a function of lagged latent variables. That is:

$$\mathbf{H}_t^{GJR} = \gamma\left[\boldsymbol{\varepsilon}_{t-1}\left(S_{t-2}\right), \mathbf{H}_{t-1}^{GJR}\right]. \tag{30}$$

Since both GARCH processes (29) and (30) are systems of first order difference equations, they have an infinite memory, that is $\mathbf{H}_t^j$ depends on $\mathbf{H}_{t-1}^j$ and through this on $\mathbf{H}_{t-2}^j$, and so on, back to the beginning of the process, for $j = Nelson, GJR$. But each of the lagged $\mathbf{H}_t^j$ terms is, in turn, a function of lagged error terms $\boldsymbol{\varepsilon}_t$ (either standardized or in level) and, indirectly, of lagged state variables S_t. Therefore, $\mathbf{H}_t^j$ will depend on the entire sequence of state variables, starting with S_{t-2}, i.e. $\mathbf{H}_t^j = \psi^j\left(S_{t-2}, S_{t-3}, \ldots, S_0\right)$. The associated likelihood function has to take into account all possible paths of the latent variables. Assume, for the sake of simplicity, that there is only one state variable and $k = 2$ regimes. At the tth observation, the likelihood function will have 2^t components, rendering estimation soon unfeasible[17]. Bekaert and Harvey (1995), in estimating a CAPM model with a regime switching ARCH-M technique, remove the path dependence by averaging out, for each period, the error terms coming from different regimes. Gray (1996) faces a similar problem in the context of a regime switching univariate GARCH model. He breaks down the path dependence of the conditional variance by weighting each of its state-dependent components with *ex ante* probabilities. Dueker (1997) keeps track of only the two most recent state-dependent variances, which are again averaged out. Following Bekaert and Harvey (1995), the error terms coming from each possible regime are averaged out. Specifically, since only the asset return equations show switching parameters, only the associated disturbances need to be averaged out. Here it is assumed that the latent variable S_t takes on two possible values. To each of these values is associated a set of asset return equations, which, in turn, will generate two disturbance vectors, $\boldsymbol{\varepsilon}_{M,t}\left(S_{t-1} = k\right)$, $k = 1, 2$. Weighting them out with the *ex ante* probabilities will remove the state-dependence:

$$\overset{\wedge}{\boldsymbol{\varepsilon}}_{M,t} = \sum_{j=1}^{2} P\left(S_{t-1} = j \left| \Im_{t-2};\ \boldsymbol{\theta}\right.\right) \boldsymbol{\varepsilon}_{M,t}\left(S_{t-1} = j\right). \tag{31}$$

$\overset{\wedge}{\boldsymbol{\varepsilon}}_{M,t}$ will then be used in each of the two GARCH processes and the matrix $\mathbf{H}_t^j$ will no longer depend on any state variable. In other words, equation (29) becomes:

$$\mathbf{H}_t^{Nelson} = f\left(\overline{\boldsymbol{\upsilon}}_{t-1}, \mathbf{H}_{t-1}^{Nelson}\right), \tag{32}$$

[17]For further details on regime switching GARCH models and the associated estimation problems see Cai (1994) and Hamilton and Susmel (1994).

where $\overline{\boldsymbol{\upsilon}}_t = \left[\overset{\wedge}{\boldsymbol{\upsilon}}_{M,t} \;\; \boldsymbol{\upsilon}_{F,t}\right]'$. $\overset{\wedge}{\boldsymbol{\upsilon}}_{M,t}=\overset{\wedge}{\boldsymbol{\varepsilon}}_{M,t} (\div) diag\sqrt{\mathbf{H}_{M,t}^{Nelson}}$, where $(\div)$ denotes the Hadamard (element by element) matrix division, while $diag\sqrt{\mathbf{H}_{M,t}^{Nelson}}$ is a $n \times 1$ vector which includes the square roots of the first n elements on the main diagonal of the $\mathbf{H}_t^{Nelson}$ matrix; similarly, $\boldsymbol{\upsilon}_{F,t} = \boldsymbol{\varepsilon}_{F,t}(\div) diag\sqrt{\mathbf{H}_{F,t}^{Nelson}}$, where $diag\sqrt{\mathbf{H}_{F,t}^{Nelson}}$ is a $m \times 1$ vector which includes the square roots of the last m elements on the main diagonal of the $\mathbf{H}_t^{Nelson}$ matrix. By the same token, equation (30) reduces to:

$$\mathbf{H}_t^{GJR} = g\left(\overline{\boldsymbol{\varepsilon}}_{t-1}, \mathbf{H}_{t-1}^{GJR}\right), \tag{33}$$

where $\overline{\boldsymbol{\varepsilon}}_t = \left[\overset{\wedge}{\boldsymbol{\varepsilon}}_{M,t} \;\; \boldsymbol{\varepsilon}_{F,t}\right]'$.

Thus, process (11) and (13) become, respectively:

$$\mathbf{H}_{t+1} = \mathbf{H}_0 \odot (\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}') + \mathbf{aa}' \odot \eta\left(\overline{\boldsymbol{\upsilon}}_t\right) \eta\left(\overline{\boldsymbol{\upsilon}}_t\right)' + \boldsymbol{\zeta\zeta}' \odot \overline{\boldsymbol{\upsilon}}_t \overline{\boldsymbol{\upsilon}}_t' + \mathbf{bb}' \odot \mathbf{H}_t, \tag{34}$$

$$\mathbf{H}_{t+1} = \mathbf{H}_0 \odot (\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}') + \mathbf{aa}' \odot \overline{\boldsymbol{\varepsilon}}_t \overline{\boldsymbol{\varepsilon}}_t' + \boldsymbol{\zeta\zeta}' \odot \overline{\boldsymbol{\eta}}_t \overline{\boldsymbol{\eta}}_t' + \mathbf{bb}' \odot \mathbf{H}_t, \tag{35}$$

where the elements contained in $\overline{\boldsymbol{\eta}}_t$ are modeled as in (14)-(17), but the residuals for the asset mean equations are set equal to $\overset{\wedge}{\boldsymbol{\varepsilon}}_{M,t}$. Thus, for instance, $\overline{\eta}_{stock,t} = -I\left(\overset{\wedge}{\varepsilon}_{stock,t} < 0\right) \overset{\wedge}{\varepsilon}_{stock,t}$.

The vector of unknown parameters $\boldsymbol{\theta}$ is estimated by combining two numerical algorithms of optimization: The Newton-Raphson and the BHHH (Berndt, Hall, Hall and Hausman, 1974) methods. The former, though more primitive, has proved useful in identifying the optimal region in the parameter space; the latter is a refinement of the first and is widely used in the empirical GARCH literature. The maximization is performed using the Constrained ML module in GAUSS software.

5 Empirical Results

In this Section Merton's (1973) Intertemporal CAPM is tested. The prices of risk are first held constant and then allowed to vary over time according to the filter proposed by Hamilton (1988, 1989, 1990, 1994). When prices are constant, the conditional second moments implied by Merton's (1973) asset pricing theory are assumed to follow first the generalized Nelson parametrization and then the generalized GJR model. The former representation turns out not to be robust to different sets of starting values and optimization algorithms. The generalized GJR model, on the other hand, proves more suited for the data set employed here. While the market risk premium plays the more relevant role in the determination of the total premium for equities, the reverse occurs for T-bills, where the market premium is negligible when compared to the intertemporal premium. Both the intertemporal and the market premia are, instead, important determinants of the bond total premium. Major financial turmoils cause an increase of the premia required to hold risky assets and spillovers from one market to another are also captured. When the prices of risk are allowed to be time-varying, the generalized Nelson model suffers from numerical problems in the optimization procedure. Therefore, second moments are estimated only with the GJR parametrization. Finally, the Markov transition probabilities are first assumed to be constant, and then allowed to evolve through time with the use of some information variables. This latter extension, however, does not seem to improve the final results.

5.1 Conditional Intertemporal CAPM with Constant Prices of Risk

Merton's (1973) asset pricing restrictions are tested considering three assets, (US stocks, 6-month Treasury bills, and 10-year government bonds), and one additional pricing factor (the growth rate of industrial production). Therefore equation (5) becomes:

$$\mathbf{R}_{t+1} = \lambda_M \sum_{j=1}^{3} \mathbf{h}_{j,t+1} w_{j,t} + \lambda_F \mathbf{h}_{4,t+1} + \boldsymbol{\varepsilon}_{M,t+1}, \tag{36}$$

where $\mathbf{R}_{t+1}$ includes excess returns on stocks, excess holding yields on the 6-month T-bills and the 10-year government bonds. The dynamic of the industrial production growth rate (see equation (6)) is assumed to be driven by itself lagged once and by the three month T-bill return lagged six times, $3MTR_{t-5}$ (source: CRSP). This specification is motivated by the fact that the IP series is serially autocorrelated while interest rates usually show their effects on investments only after some time:

$$IP_{t+1} = k_0 + k_1 IP_t + k_2 3MTR_{t-5} + \varepsilon_{IP,t+1}. \tag{37}$$

The disturbance vector, $\boldsymbol{\varepsilon}_{t+1} = [\boldsymbol{\varepsilon}_{M,t+1} \ \varepsilon_{IP,t+1}]'$, is supposed to be conditionally Student t-distributed:

$$\boldsymbol{\varepsilon}_{t+1} \left| \Im_t \right. \sim Student - t \ \left(0, \mathbf{H}_{t+1}, \nu\right),$$

where $\mathbf{H}_{t+1}$ is a 4×4 asymmetric GARCH process. It is first assumed that $\mathbf{H}_{t+1}$ follows the generalized Nelson parametrization described by (11) and next the generalized GJR model of equation (13).

The vector of parameters $\boldsymbol{\theta}$ is estimated with the log likelihood function (27), where the density function $\phi\left(\mathbf{G}_{t+1}\left|\Im_t;\boldsymbol{\theta}\right.\right)$ becomes a state-independent multivariate Student t-distribution:

$$\phi\left(\mathbf{G}_{t+1}\left|\Im_t;\boldsymbol{\theta}\right.\right) =$$

$$= \frac{\Gamma\left(\frac{\nu+g}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left[\pi\left(\nu-2\right)\right]^{\frac{g}{2}}\left|\mathbf{H}_{t+1}\left(\boldsymbol{\theta}\right)\right|^{\frac{1}{2}}}\left[1+\frac{\boldsymbol{\varepsilon}_{t+1}\left(\boldsymbol{\theta}\right)'\mathbf{H}_{t+1}\left(\boldsymbol{\theta}\right)^{-1}\boldsymbol{\varepsilon}_{t+1}\left(\boldsymbol{\theta}\right)}{\nu-2}\right]^{-\frac{\nu+g}{2}}, \tag{38}$$

where $g=4$, and $\mathbf{G}_{t+1}=\left[\mathbf{R}_{t+1}\ IP_{t+1}\right]'$[18].

5.1.1 The Generalized Nelson Model

In this Section the conditional covariance matrix of the error term vector is assumed to follow the generalized Nelson model (11). Results are reported in Table 3. The prices of market and intertemporal risk are both significant and of the expected sign, while only a_1 and a_2 are significantly different from zero. b_4 is larger than one, which makes the GARCH process non-stationary (see Zakoian, 1994). However, even if the mathematical properties for stability are not fulfilled, depriving the model of a long-term solution, estimations still bring some economic intuition. Interestingly, ζ_1 is negative and significant, showing that negative past innovations have a larger impact than positive shocks of the same magnitude on the conditional stock variance. The opposite occurs for T-bills, with ζ_2 larger than zero (although not significant). Notice that the signs found for ζ_1 and ζ_2 are consistent with what is expected on the grounds of the Engle and Ng's (1993) tests (see Section 3). The intuition is therefore the same. The fact that $\zeta_1 < 0$ indicates that both the leverage effect and the feedback volatility hypothesis are at work: When an unanticipated negative shock occurs, the debt-to-equity ratio increases, making stocks a riskier asset. If the rise in volatility is anticipated, investors will require higher returns, generating a fall in equity prices which causes return shocks. These negative innovations coupled with volatility persistence explain why the conditional variance increases. The leverage effect does not play any role for the T-bills. Contrary to Campbell and Hentschell's (1992) theory, but consistent with the Engle and Ng's (1993) test, more weight is attached to past positive than to negative innovations. As for government

[18] Estimations have first been carried out assuming that conditional residuals are normally distributed. When the Student $t-$distribution is considered, parameters exhibit the same sign and about the same size as those obtained under normality. The use of a Student$-t$ distribution has, therefore, given rise to efficiency gains (see Gouriéroux and Monfort, 1995, for further details).

bonds and the growth rate of industrial production, the generalized Nelson model cannot accommodate for the fact that their volatility increases when both positive and negative past innovations occur. Therefore the sign attached to ζ_3 and ζ_4 (which turn out to not even be significant) does not bring much intuition. Overall, the generalized Nelson model seems to be suited for portfolios of assets and factors whose conditional volatility, say, either goes up or down after, respectively, a negative or positive past shock. Portfolios of equities only, for instance, might be good candidates for this multivariate asymmetric GARCH parametrization.

The parameters involved in the specification of the state variable are all significant and of the expected sign. In particular, k_1 is positive, showing that the series has a certain memory, while k_2 is negative, indicating that a change in interest rates has an opposite effect on the industrial production growth rate.

Unfortunately, the optimum that has been reached, besides being non-stationary, is not robust to different sets of starting values or algorithms other than Newton-Rapson, which questions the reliability of the parameters that have been found.

5.1.2 The Generalized GJR Model

In this Section the conditional second moments of Merton's (1973) asset pricing model are assumed to follow the asymmetric GARCH process given by (13), where the four elements included in $\boldsymbol{\eta}_t$ are described by equations (14) to (17). Results are shown in Table 4, Panel A. The two prices of risk, λ_M and λ_F, are, respectively, positive and negative, in line with the theory, as well as significantly different from zero. All the parameters implied by the GARCH process (13), apart from ζ_{IP1}, as well as those governing the dynamic of the state variable, are significant. To test the asymmetric against the symmetric GARCH process as well as whether some of the ζ's parameters are equal, several Wald tests have been carried out (See Table 4, Panel B). The two null hypotheses that $\zeta_{bond1} = \zeta_{bond2}$ and $\zeta_{IP1} = \zeta_{IP2}$ are rejected at any conventional level, as well as the null that all the six ζ's are jointly equal to zero. Furthermore, the optimum that has been reached is now robust to different sets of starting values and to two optimization algorithms (Newton-Rapson and BHHH).

In Figure 1 the market (MRP), intertemporal (IRP), and total (TRP) risk premia for each asset are plotted. The formulae to compute each premium are given by equation (39), (40), and (41), respectively:

Table 3
Estimation results for the Merton's (1973) Intertemporal CAPM with constant prices of risk
Conditional second moments specification: Generalized Nelson model

The estimated model is:

$$\mathbf{R}_{t+1} = \lambda_M \sum_{j=1}^{3} \mathbf{h}_{j,t+1} w_{j,t} + \lambda_F \mathbf{h}_{4,t+1} + \boldsymbol{\varepsilon}_{M,t+1},$$

$$IP_{t+1} = k_0 + k_1 IP_t + k_2 3MTR_{t-5} + \varepsilon_{IP,t+1},$$

The error terms are conditionally Student t-distributed, i.e. $\boldsymbol{\varepsilon}_{t+1} | \Im_t \sim Student - t \; (0, \mathbf{H}_{t+1}, \nu)$. The conditional covariance matrix follows the generalized Nelson process described by equation (11):

$$\mathbf{H}_{t+1} = \mathbf{H}_0 \odot (\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}') + \mathbf{aa}' \odot \eta(\boldsymbol{\upsilon}_t) \, \eta(\boldsymbol{\upsilon}_t)' + \boldsymbol{\zeta\zeta}' \odot \boldsymbol{\upsilon}_t \boldsymbol{\upsilon}_t' + \mathbf{bb}' \odot \mathbf{H}_t.$$

where $\eta(\boldsymbol{\upsilon}_t) = |\boldsymbol{\upsilon}_t| - E|\boldsymbol{\upsilon}_t|$. $\boldsymbol{\upsilon}_t = \boldsymbol{\varepsilon}_t (\div) diag\sqrt{\mathbf{H}_t}$, where $(\div)$ denotes the Hadamard (element by element) matrix division, while $diag\sqrt{\mathbf{H}_t}$ is a vector which includes the square roots of the elements on the main diagonal of the $\mathbf{H}_t$ matrix. Since the error terms are assumed to be Student t-distributed, $E|\boldsymbol{\upsilon}_t| = \left\{ (2/\pi)^{1/2} [\Gamma(\nu - 1)/2] / \Gamma(\nu/2) \right\} \mathbf{1}$.

λ_M	λ_F	k_0	k_1	k_2	ν
0.046	−1.434	0.470	0.228	−0.534	4.005
(0.012)	(0.210)	(0.057)	(0.043)	(0.104)	(0.135)

	R^{stock}	R^{bill}	R^{bond}	IP
a_i	−0.088	−0.244	−0.092	−0.018
	(0.040)	(0.033)	(0.051)	(0.019)
ζ_i	−1.869	0.009	0.063	−0.030
	(0.604)	(0.017)	(0.054)	(0.008)
b_i	0.907	0.960	0.920	1.003
	(0.042)	(0.011)	(0.065)	(0.001)

Log-likelihood function	− 7.570

Standard errors, based on the matrix of second derivatives of the log likelihood function, are shown in parentheses.

Table 4
Estimation results for the Merton's (1973) Intertemporal CAPM with constant prices of risk
Conditional second moments specification: Generalized GJR model

The estimated model is:

$$\mathbf{R}_{t+1} = \lambda_M \sum_{j=1}^{3} \mathbf{h}_{j,t+1} w_{j,t} + \lambda_F \mathbf{h}_{4,t+1} + \boldsymbol{\varepsilon}_{M,t+1},$$

$$IP_{t+1} = k_0 + k_1 IP_t + k_2 3MTR_{t-5} + \varepsilon_{IP,t+1},$$

The error terms are conditionally Student t-distributed, i.e. $\boldsymbol{\varepsilon}_{t+1} | \Im_t \sim Student - t\ (0, \mathbf{H}_{t+1}, \nu)$. The conditional covariance matrix follows the generalized GJR process described by equation (13):

$$\mathbf{H}_{t+1} = \mathbf{H}_0 \odot (\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}') + \mathbf{aa}' \odot \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' + \boldsymbol{\zeta}\boldsymbol{\zeta}' \odot \boldsymbol{\eta}_t \boldsymbol{\eta}_t' + \mathbf{bb}' \odot \mathbf{H}_t.$$

where $\eta_{stock,t} = -I(\varepsilon_{stock,t} < 0)\,\varepsilon_{stock,t},\ \eta_{bill,t} = I(\varepsilon_{bill,t} > 0)\,\varepsilon_{bill,t},\ \eta_{bond,t} = \begin{cases} \zeta_{bond1} \cdot \varepsilon_{bond,t} & \text{if } \varepsilon_{bond,t} > 0 \\ -\zeta_{bond2} \cdot \varepsilon_{bond,t} & \text{if } \varepsilon_{bond,t} \leq 0 \end{cases}$, and $\eta_{IP,t} = \begin{cases} \zeta_{IP1} \cdot \varepsilon_{IP,t} & \text{if } \varepsilon_{IP,t} > 0 \\ -\zeta_{IP2} \cdot \varepsilon_{IP,t} & \text{if } \varepsilon_{IP,t} \leq 0 \end{cases}$.

Panel 4A: Estimation results

λ_M	λ_F	k_0	k_1	k_2	ν
0.037	−0.524	0.548	0.191	−0.453	4.637
(0.011)	(0.119)	(0.081)	(0.048)	(0.144)	(0.490)

	R^{stock}	R^{bill}	R^{bond}	IP
a_i	0.308	0.354	0.350	0.341
	(0.030)	(0.033)	(0.024)	(0.030)
ζ_i	0.297	0.195		
	(0.058)	(0.056)		
b_i	0.932	0.931	0.936	0.940
	(0.014)	(0.013)	(0.009)	(0.011)

ζ_{bond1}	ζ_{bond2}	ζ_{IP1}	ζ_{IP2}
−0.198	0.348	0.085	−0.323
(0.054)	(0.053)	(0.065)	(0.057)

Log-likelihood function	– 7.377

Standard errors, based on the matrix of second derivatives of the log likelihood function, are shown in parentheses.

Panel 4B: Specification tests for the parameters governing the asymmetric effects

Null Hypothesis	D. of F.	χ^2_m	p-value
Do positive and negative past shocks have the same effect on government bond volatility? $H_0 : \zeta_{bond1} = \zeta_{bond2}$	1	47.463	0.000
Do positive and negative past shocks have the same effect on industrial production growth rate volatility? $H_0 : \zeta_{IP1} = \zeta_{IP2}$	1	21.260	0.000
Are all the parameters governing the asymmetric effect in conditional second moments significant? $H_0 : \zeta_{stock} = \zeta_{bill} = \zeta_{bond1} = \zeta_{bond2} = \zeta_{IP1} = \zeta_{IP2} = 0$	6	90.775	0.000

The Wald test statistics is asymptotically distributed as χ^2_m, with m indicating the degrees of freedom.

$$MRP_{i,t+1} = \lambda_M \sum_{j=1}^{3} Cov\left(R_{i,t+1}, R_{j,t+1} \left|\Im_t\right.\right) w_{j,t}, \qquad i = stock, bill, bond. \tag{39}$$

$$IRP_{i,t+1} = \lambda_F \left(CovR_{i,t+1}, IP_{t+1} \left|\Im_t\right.\right), \qquad i = stock, bill, bond. \tag{40}$$

$$TRP_{i,t+1} = MRP_{i,t+1} + IRP_{i,t+1}, \qquad i = stock, bill, bond. \tag{41}$$

First, notice that the price of intertemporal risk is negative. This means that the elasticity of marginal utility of wealth with respect to the growth rate of industrial production is positive, i.e. $J_{WIP} > 0$. When IP goes up, signaling an expansion in the economy, the marginal utility of wealth increases as well. Therefore, the sign of the intertemporal premium will depend on the sign of the covariance between each asset and the growth rate of industrial production. If this covariance is positive, the total premium investors require for each security will decrease; conversely, when the covariance is negative, the total premium will increase.

As far as stocks are concerned, market premia are seen to increase in two recessional periods as reported by NBER, namely in the early '70s, and during the contraction caused by the first oil shock. The raise in premia is much more modest in the first half of the '80s, and in the early '90s, again

Figure 1: Components of excess returns

Conditional Intertemporal CAPM with constant prices of risk. Conditional second moment specification: Generalized GJR model. The scale along the vertical axes is in units of percent.

Figure 1A: Stock excess returns

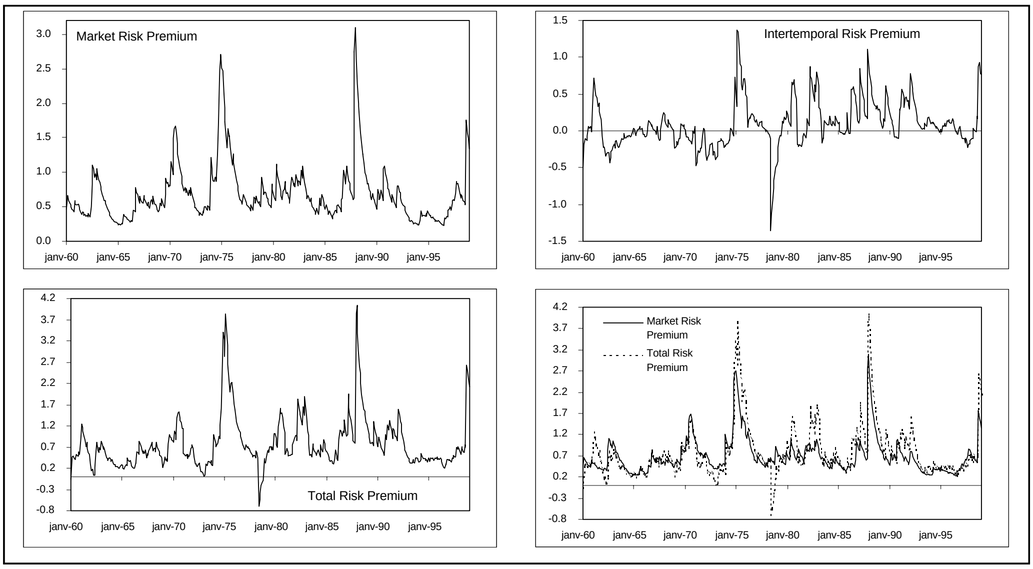


Figure 1B: Excess holding yields on 6-month T-bills

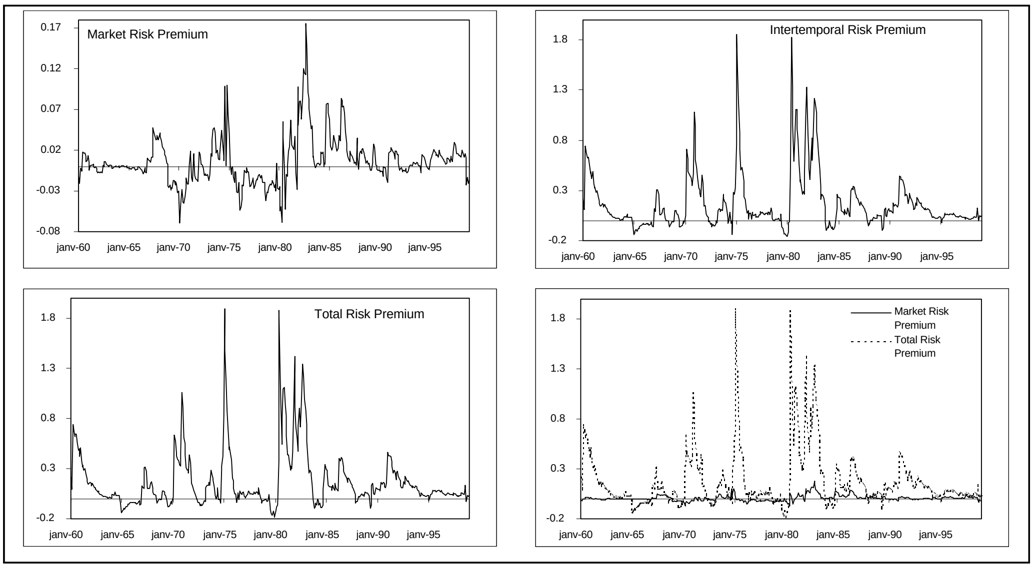
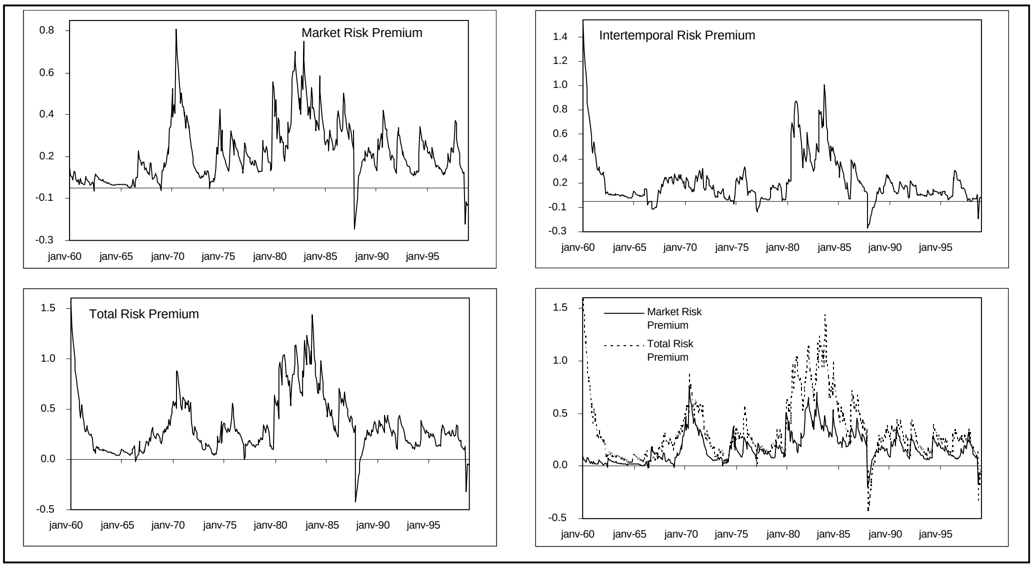


Figure 1C: Excess holding yields on 10-year government bonds



periods of contraction according to the NBER. Furthermore, market premia surge during the 1987 stock market crash, and finally during the 1997/98 Asian-Russian-Latin American crises. The intertemporal risk premium is almost always negative until the middle of the '70s and thereafter almost always positive, with the exception of the period following the first oil shock. Such a time evolution can have an appealing interpretation: In the middle of the '70s, when the consequences of the oil shock hit the American economy heavily, the intertemporal risk premium shows its highest peak, increasing the total premium investors require to hold equities. In such a contraction phase, stocks were not good hedges against business cycle downturns. On the contrary, the deepest through which occurred in the second half of the '70s might indicate the willingness to hold stocks, due to the recovery of the economy. Overall, from the '80s on, the intertemporal risk premium is positive, reflecting a negative correlation between stock returns and the growth rate of industrial production. Such a negative conditional covariance can be explained considering that the stock market and growth rate of industrial production are linked but are not in phase (see, for instance, Hamilton and Lin, 1996, Chauvet and Potter, 1998, and Chauvet, 1999)[19]. In this case, the total premium investors require to hold equities increases. Finally, notice that the market premium constitutes the main component of the total premium.

The largest movements in the T-bill market risk premia occurre after the first oil shock and between 1979 and 1982. During this latter period the

[19]Interestingly, also the unconditional covariance between the two time series is negative and equal to -0.154.

Federal Reserve (Fed) changed its monetary policy in that, instead of targeting interest rates, it preferred to control nonborrowed reserves. Interest rates went up, next down, and increased again afterwards. Volatility, thus, surged for government fixed income securities. As a consequence the market premium is seen first to decrease and then to rise. The money market also suffered from the stock market crash, as evidenced by the peak at the end of 1987. The intertemporal risk premium is positive most of the time, indicating that the covariance between returns on T-bills and the growth rate of industrial production is almost always negative, which finds its rationale in the negative correlation between interest rates and investments. Indeed, it reaches its highest peaks at the beginning and in the middle of the '70s and in the first half of the '80s. This may reflect the fact that T-bills are not considered a good hedge against "bad times". A result which is intuitive, if one considers that returns on T-bills show a sort of pro-cyclical pattern: Usually central banks lower (increase) interest rates at business cycle troughs (peaks), in an attempt to stimulate (cool down) the economy. Overall, the total premium that comes out is much larger than the market risk premium: When the two premia are plotted on the same graph, the latter is negligible when compared to the former, indicating that the total premium is almost entirely driven by the intertemporal premium.

As far as government bonds are concerned, the market risk premium increases in the early and middle '70s, periods of business cycle downturns, as well as between 1979 and 1982, again due to the change in the Fed's monetary policy. The market risk premium plunges in correspondence of the October 1987 stock market crash: Probably investors feared a weakening in the economy and expected that bonds would have performed better than shares[20]. Indeed, growing numbers of institutional investors, who traditionally used to hold mainly equities, switched to bonds. The decrease in the bond market premium reflects this post-crash "flight to quality". A similar phenomenon occurred at the end of 1998, when the market premium is again negative. The intertemporal risk premium is positive most of the time, indicating that the conditional covariance between returns on bonds and the growth rate of industrial production is almost always negative. Government bonds, therefore, cannot be considered good hedges against business cycle downturns. Finally, once more, the intertemporal premium constitutes an important portion of the total.

All in all, if the "true" model is Merton's (1973) Intertemporal CAPM but asset pricing and portfolio management are carried out with the tradi-

[20]In another version of the generalized GJR model, where the same weight is attached to both positive and negative past bond and IP shocks, the bond market risk premium is seen to increase after October 1987. This indicates that the specification proposed here is preferable.

tional static CAPM, where the total and the market premia coincide, then investors might be seriously mistaken. However, these results must be interpreted with caution. Even if Merton's model is more satisfactory than the traditional one, it provides little guidance for the choice of the additional pricing factors as well as their number, a topic worth a separate study (see, for instance, Chen, Roll, and Ross, 1986, and Fama, 1998).

5.1.3 The Asymmetric Effects in Conditional Second Moments: Graphical Evidence and Robust Test Statistics

This Section provides graphical evidence and diagnostic robust test statistics to show that the generalized GJR model (13) - (17) is able to capture the asymmetric responses of variances and covariances to past shocks. In particular, news impact functions and surfaces (see Engle and Ng, 1993, and Kroner and Ng, 1998 for details) are plotted. Furthermore, the robust conditional moment test statistics proposed by Kroner and Ng (1998) are computed.

Notice that, due to the use of a diagonal GARCH process, by construction, the (i, j)th element of the covariance matrix only depends on the corresponding past (i, j)th innovation, while past cross products of the ε_t's do not enter into second moment specification. Therefore, here news impact curves establish a relationship between the current conditional volatility of the variable i and its own last period shocks (news) $\varepsilon_{i,t}$, evaluating lagged conditional variances at their unconditional sample mean level. Similarly, news impact surfaces plot conditional covariances between variables i and j against past innovations $\varepsilon_{i,t}\varepsilon_{j,t}$, holding lagged conditional covariances constant at their unconditional value. In Figure 2, such curves and surfaces are represented for excess returns on stocks, T-bills, and government bonds, as well as for the growth rate of industrial production.

As far as equities are concerned, negative shocks increase the conditional variance much more than positive innovations of the same magnitude, while the opposite occurs for T-bills. Like equity variance, conditional volatility of bonds as well as that of the growth rate of industrial production respond more to past negative than positive shocks, though the difference is not that large.

To better see the effect of the asymmetric component in the generalized GJR model, Figure 2B plots the conditional covariance between two generic variables i and j when the term $\boldsymbol{\zeta}\boldsymbol{\zeta}' \odot \boldsymbol{\eta}_t\boldsymbol{\eta}_t'$ is set equal to zero and the asymmetric GARCH process (13) reduces to the symmetric specification

Figure 2: News impact curves and surfaces of the generalized GJR model

The scale along the vertical axes is in units of percent.

Figure 2A: News impact curves

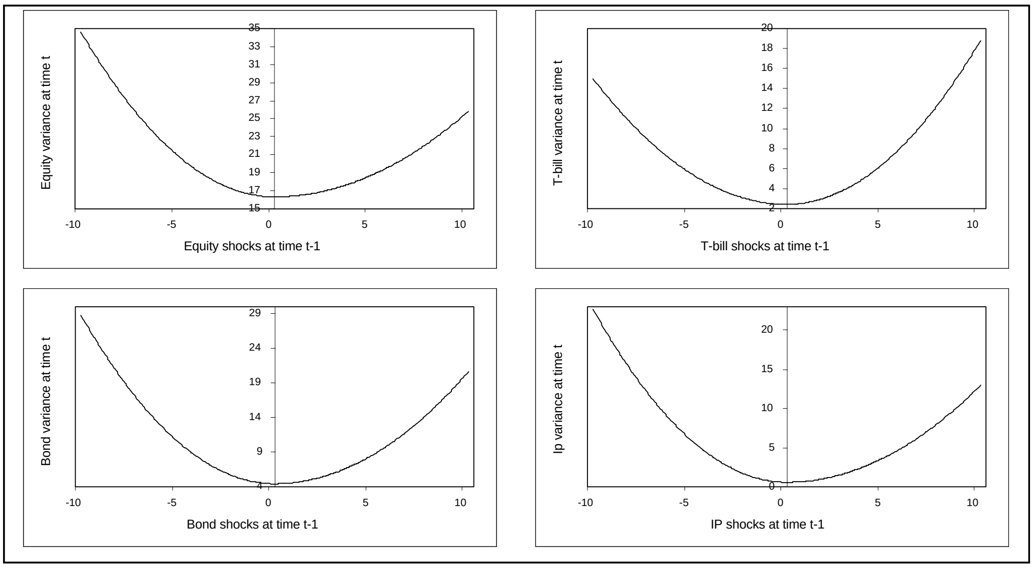


Figure 2B: News impact surfaces: Symmetric specification

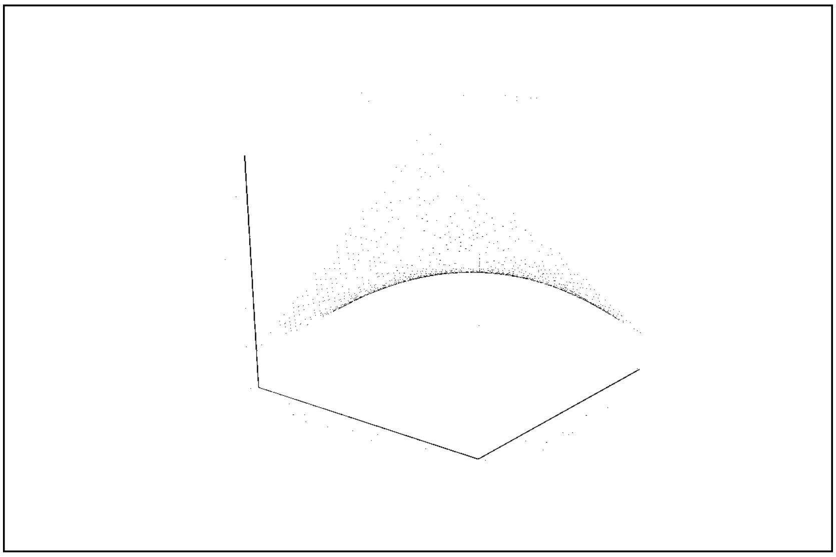
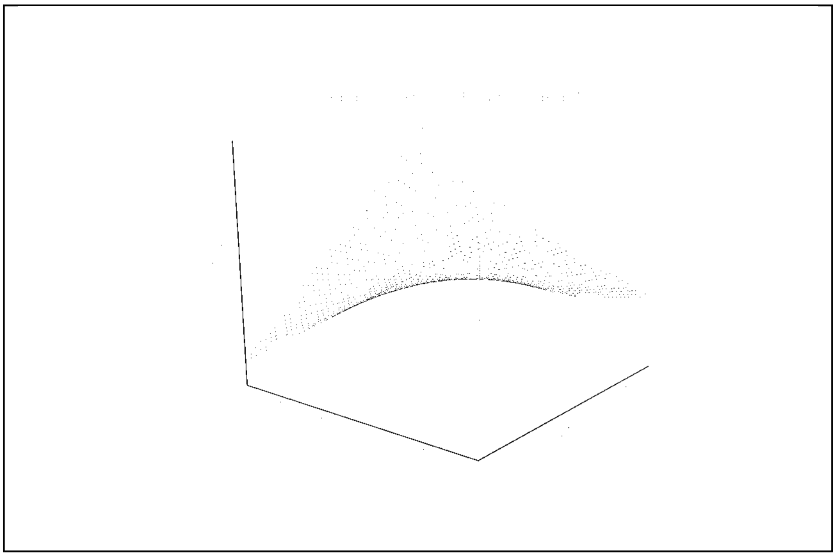


Figure 2C: News impact surfaces: Asymmetric specification



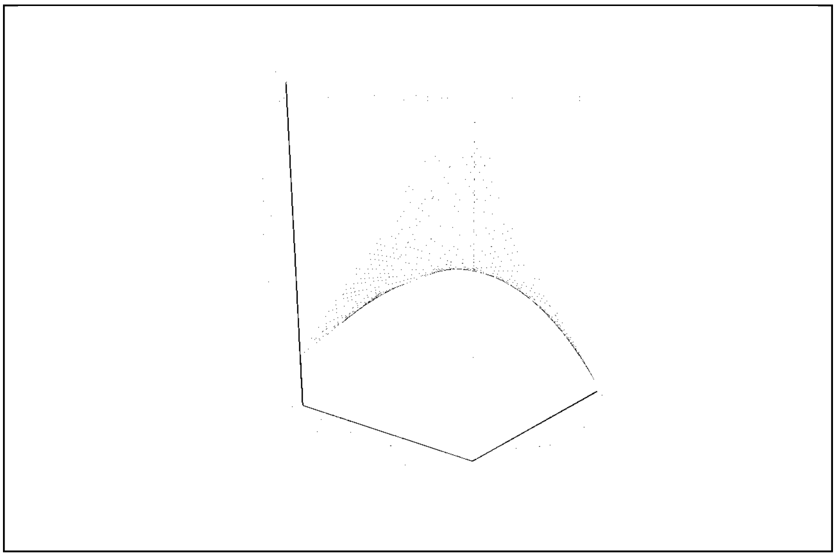
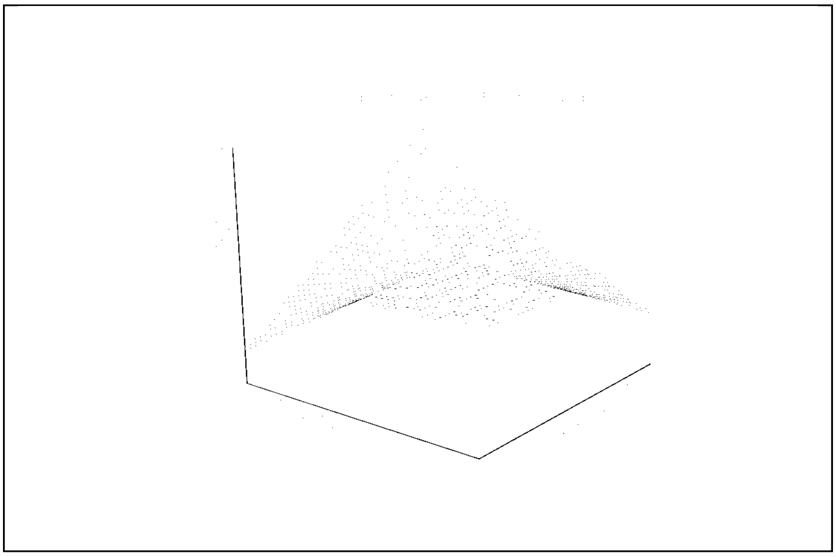


Figure 2C (contd.): News impact surfaces: Asymmetric specification



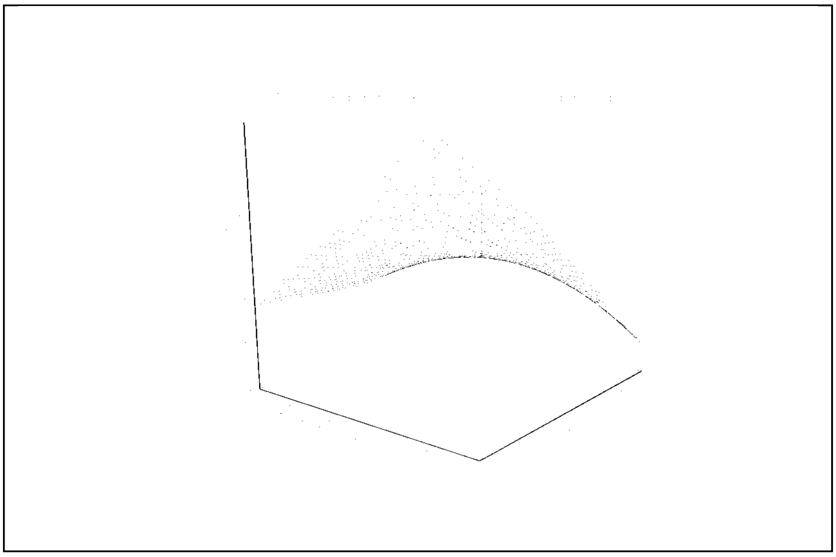
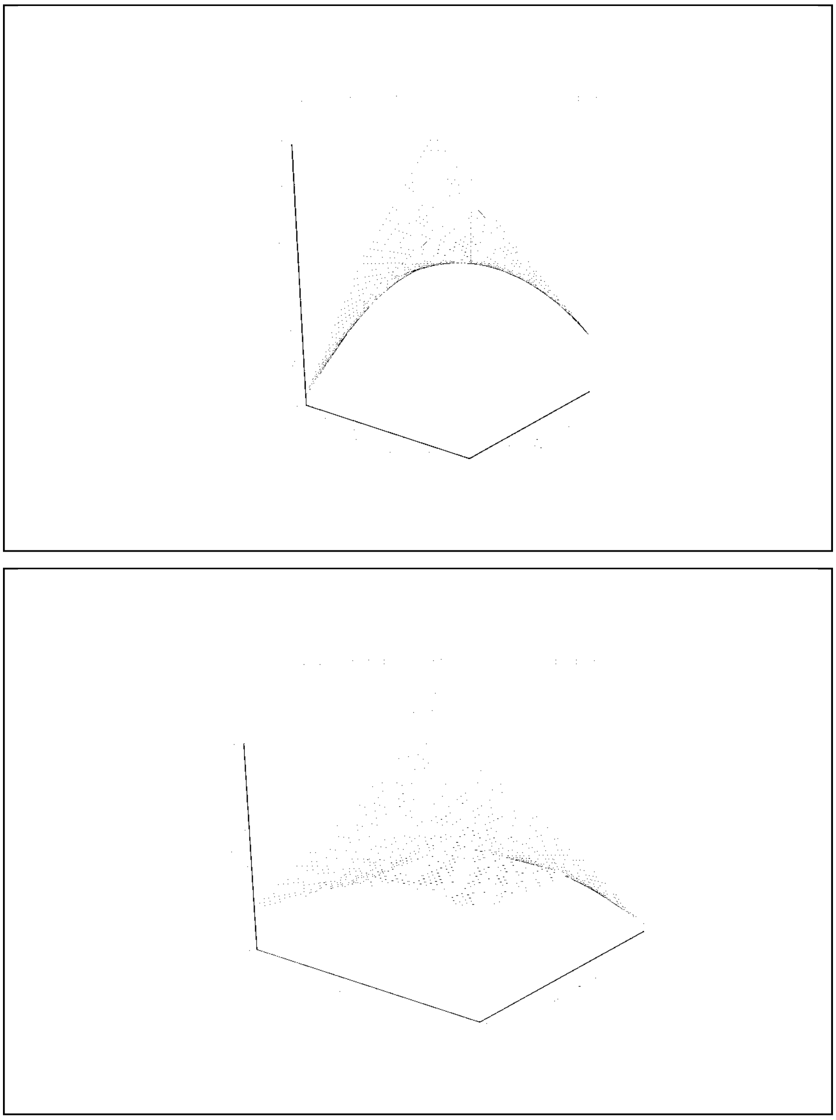


Figure 2C (contd.): News impact surfaces: Asymmetric specification



given by equation (9)[21]. In this case, when past shocks show the same sign and both take on extreme values, the conditional covariance reaches its highest (same) peaks. Conversely, when extreme past innovation values of different signs are combined, the conditional covariance is seen to be at its lowest (same) levels. Such a representation is used as a benchmark. In Figures 2C, instead, second moments obtained from estimating model (13) are plotted. Large positive equity shocks together with small negative T-bill innovations increase the conditional covariance, whereas the remaining part of the surface does not change much if compared to Figure 2B. Similarly, the conditional covariance between stocks and bonds responds asymmetrically to large positive equity shocks and small bond innovations. The conditional covariance between equities and IP is seen to decrease when bad news hits the stock market while there are negative innovations for the state variable. Good news to the bond market and bad news to the T-bill market bring the conditional covariance up. When negative T-bill shocks are associated with positive IP shocks, the correspondent conditional covariance decreases. Finally, the surface representing the conditional covariance between bonds and IP exhibit a shape similar to that of stocks and the industrial production growth rate.

Diagnostic statistics for the standardized residuals in level, i.e. $\varepsilon^{*}_{i,t+1} = \varepsilon_{i,t+1} h^{-1/2}_{ii,t+1}$, and squared, i.e. $\varepsilon^{*2}_{i,t+1} = \varepsilon^{2}_{i,t+1} h^{-1}_{ii,t+1}$, of the generalized GJR model are presented in Table 5. The Ljung-Box portmanteau statistics does not reject the null of no simultaneous autocorrelation for stocks, bonds, and the industrial production growth rate, but it does for T-bills, indicating that an autoregressive term in the T-bill mean equation might provide a better fit.

The Ljung-Box tests, however, are not suited to check whether a GARCH process is able to capture the asymmetric responses of conditional second moments to past innovations. Therefore, to check if the generalized GJR model is mis-specified, the robust conditional moment tests of Kroner and Ng (1998) are applied (see the authors for further details). These tests are based on mis-specification indicators, x_g. In particular, one has to compute $x_l = I\left(\varepsilon_{i,t} < 0\right)$, $l = i = 1, \ldots, 4$; $x_{1m} = I\left(\varepsilon_{i,t} < 0; \varepsilon_{j,t} < 0\right)$, $x_{2m} = I\left(\varepsilon_{i,t} < 0; \varepsilon_{j,t} > 0\right)$, $x_{3m} = I\left(\varepsilon_{i,t} > 0; \varepsilon_{j,t} < 0\right)$, $x_{4m} = I\left(\varepsilon_{i,t} > 0; \varepsilon_{j,t} > 0\right)$, $i, j = 1, \ldots, 4$, for a total of 24 combinations; $x_n = \varepsilon^2_{i,t} I\left(\varepsilon_{j,t} < 0\right)$, $i, j = 1, \ldots, 4$, for a total of 16 combinations. Finally, for the sake of completeness, another set of mis-specification indicators can be proposed, namely $x_p = \varepsilon^2_{i,t} I\left(\varepsilon_{j,t} > 0\right)$, $i, j = 1, \ldots, 4$, for a total of 16 combinations. The last set of indicators can be seen as an extension to a multivariate context of the

[21] When plotting news impact curves and surfaces, since lagged conditional (co)variances are evaluated at their unconditional sample mean level, they are mere scaling factors. For the sake of simplicity, when $\boldsymbol{\zeta}\boldsymbol{\zeta}' \odot \boldsymbol{\eta}_t\boldsymbol{\eta}_t' = \mathbf{0}$, they are set equal to one.

Table 5
Diagnostic tests for the standardized residuals of the generalized GJR model

	R^{stock}	R^{bill}	R^{bond}	IP
Mean	-0.048	0.233	-0.082	-0.109
Std. Dev.	0.894	0.921	0.822	0.822
L-B$_{12}$ $\left(\varepsilon^{*}_{i,t+1}\right)$	7.719	29.047^{**}	15.734	13.460
L-B$_{12}$ $\left(\varepsilon^{*2}_{i,t+1}\right)$	6.524	38.394^{**}	10.424	14.989

** denotes 1% significance level.
The Ljung-Box$_m$ (L-B$_m$) statistics tests the null hypothesis that all autocorrelation coefficients are simultaneously equal to zero up to m lags; it is asymptotically distributed as χ^2_m. It has been chosen $m = 12$ for which the critical values at 95% and 99% confidence level are 21.026 and 26.217, respectively.

positive size-bias tests introduced by Engle and Ng (1993). As before, $I(\cdot)$ represents the indicator function. Since all the possible combinations give rise to 60 figures, the test statistics are not reported here[22]. Importantly, the generalized GJR model is rejected only three times out of sixty at 1% significance level, which indicates that the representation is well specified.

5.2 Conditional Intertemporal CAPM with Time-Varying Prices of Risk

There is large consensus on the fact that both expected excess returns and conditional (co)variances change through time (see the extensive surveys of Bollerslev, Chou, and Kroner, 1992, Bera and Higgins, 1993, and Bollerslev, Engle, and Nelson, 1994). However, first and second conditional moments do not move in a one-to-one proportion. Therefore, in CAPM-type models the risk-return relationship is not constant over time. The estimation of the coefficient of risk aversion, though, is one of the "unsolved" issues in empirical finance. Being related to the investors' utility function (see Section 2), which is not observable *per se*, its evolution over time is often based on some presumptions. Nevertheless, its sign and magnitude are crucial, since they directly affect risk premia.

Although in French, Schwert, and Stambaugh (1987) the coefficient of risk aversion is positive but not significant, interestingly, it turns out to be very unstable across sample periods. In their estimations, the risk-return relation ranges from 1.510 to 7.220, when the NYSE monthly value-weighted

[22] Figures relative to the robust conditional moment tests are available from the author upon request.

index is used, and from 0.598 to 7.809, when the S&P's daily composite index is adopted. The two subperiods considered are 1928-1952 and 1953-1984, respectively. Thereafter, empirical research has continued to find support for a time-varying price of market risk: See, for instance, Harvey (1989, 1991), Chou, Engle and Kane (1992), Bekaert and Harvey (1995), Dumas and Solnik (1995), Cappiello (1998), De Santis and Gerard (1997, 1998a, 1998b), Chauvet and Potter (1998), De Santis, Gerard, and Hillion (1998), Cappiello (1999), and Bekaert and Wu (2000). All these studies, apart from Chou *et al.,* Cappiello (1999), and Chauvet and Potter, model the coefficient of risk aversion as a function of a certain number of instruments, which are designed to capture expectations in business cycle fluctuations[23]. This research documents that the price of risk increases during economic troughs while decreases during expansionary phases of the business cycle (a pattern particularly clear in Harvey, 1989, and Bekaert and Harvey, 1995). A critique that naturally arises is that the coefficient of risk aversion might follow a counter-cyclical path because the instruments through which it is modeled mimic the business cycle phases. It is true that since the parameters of the risk-return function come from an estimation procedure, the coefficient of risk aversion could, instead, well display a pro-cyclical pattern or even be constant over time. However, one can undoubtedly argue that the relationship between the price of risk and economic fluctuations is superimposed by the econometrician by his particular choice of information variables. Chou *et al.* estimate a time-varying coefficient of risk aversion first with rolling regressions and then with a Kalman filter conditional on the parameter values. Interestingly, with Chou *et al.*'s approach the time-varying risk-return relationship does not seem to be related to the business cycle. In particular, the price of market risk is high for the mid-'50s through the early '70s and low in the depression and war years as well as in the '80s. A non-linear risk-return relationship is estimated by Chauvet and Potter by combining a non-linear discrete version of the Kalman filter with Hamilton's (1988, 1989, 1990, 1994) regime switching model. The risk-return relationship turns out to be positive around business cycle troughs and negative around economic peaks. Cappiello, who employs Hamilton's filter to estimate the price of market (and intertemporal) risk, identifies two states which may reflect a switch in investors' preferences whose degree of risk aversion increases in correspondence to financial turmoil.

With few exceptions (see, for instance, De Santis and Gerard, 1998b, and

[23]The information variables widely used when modelling the price of market risk are: i) a constant term; ii) the dividend yield on the national equity index in excess of the risk-free rate; iii) the change in the 1- or 3-month T-bill rate; iv) the default premium (Moody's Baa minus Aaa bond yields); and v) the term structure spread (10-year government bond minus 3-month T-bill yield). Sometimes a January dummy is also included.

Cappiello, 1999), the literature is silent about the possibility that also the price of intertemporal risk varies over time. However, if it is true that the risk-return relation changes because both first and second moments vary over time not in a one-to-one proportion, then for the same reason the intertemporal price of risk has to be time-varying.

As seen in Section 2, both λ_M and λ_F depend on first and second derivatives of the derived utility function of wealth, which represents investors' preferences and, as such, has an unknown functional form. λ_M is also function of the wealth value, W. λ_M and λ_F are harmonic means of the prices of risk of each investor. In this paper, following Cappiello (1999), any *a priori* assumption about the functional form of the prices of risk is avoided and both λ_M and λ_F are estimated through the regime switching model of Hamilton (1988, 1989, 1990, 1994). This methodology has the advantage of making both λ_M and λ_F time-varying and letting the data "speak for themselves". As discussed in Section 4, it is assumed that the prices of risk are driven by the same two-state Markov latent variable.

When λ_M and λ_F are time-varying, the asset pricing restrictions are described by equation (22), where now $n = 3$.

The dynamic of the additional pricing factor, the growth rate of industrial production, is again given by equation (37). As before, the disturbance vector, $\boldsymbol{\varepsilon}_{t+1} = [\boldsymbol{\varepsilon}_{M,t+1}\ \varepsilon_{IP,t+1}]'$, is assumed to be conditionally Student t-distributed

$$\boldsymbol{\varepsilon}_{t+1} |\Im_t \sim Student - t\ \ (0, \mathbf{H}_{t+1}, \nu),$$

where $\mathbf{H}_{t+1}$ is a 4×4 asymmetric GARCH process. Once more, the conditional second moments implied by Merton's (1973) Intertemporal CAPM are estimated first with the generalized Nelson parametrization and next with the generalized GJR model. The vector of unknown parameters, $\boldsymbol{\theta}$, is estimated by maximizing the likelihood function (27), where the state-dependent conditional density functions are given by equation (26), with $\mathbf{G}_t = [\mathbf{R}_t\ IP_t]'$ and $g = 4$[24]. When the generalized Nelson process is employed, however, the covariance matrix of the parameters fails to invert, with the consequence that the standard errors are not available. Since this can be interpreted as a mis-specification of the model, the generalized Nelson process is discarded from the analysis. Once prices are allowed to switch, though, the generalized GJR model also shows difficulties in convergence. Therefore, to reduce the number of unknown parameters, a simplified version of the latter is estimated. In particular, the third and fourth elements of the vector $\boldsymbol{\eta}_t$ in equation (13) are modeled as follows:

[24]See footnote no. 16.

$$\eta_{bond,t} = \begin{cases} 1 \cdot \varepsilon_{bond,t} & \text{if } \varepsilon_{bond,t} > 0 \\ -1 \cdot \varepsilon_{bond,t} & \text{if } \varepsilon_{bond,t} \leq 0 \end{cases}, \tag{16'}$$

$$\eta_{IP,t} = \begin{cases} 1 \cdot \varepsilon_{IP,t} & \text{if } \varepsilon_{IP,t} > 0 \\ -1 \cdot \varepsilon_{IP,t} & \text{if } \varepsilon_{IP,t} \leq 0 \end{cases}, \tag{17'}$$

or, in terms of the indicator function:

$$\eta_{i,t} = [2 \cdot I(\varepsilon_{i,t} > 0) - 1]\, \varepsilon_{i,t} \qquad \text{for } i = bond, IP.$$

As a consequence, the third and fourth elements of the vector $\boldsymbol{\zeta}$ in equation (13) are no longer constrained to be equal to one and are estimated as free parameters. This representation assumes that the conditional volatility of government bonds and industrial production does not respond differently to positive and negative past shocks. Indeed, the parametrization in (16') and (17') amounts to just considering the absolute value of $\varepsilon_{bond,t}$ and $\varepsilon_{IP,t}$, with weights ζ_{bond} and ζ_{IP}, respectively.

Furthermore, the Markov transition probabilities implied by the use of the regime switching model are first held constant, and then allowed to vary over time.

Next subsection reports the results that are obtained.

5.2.1 The Generalized GJR Model

The estimation results for the generalized GJR model are reported in Table 6. The price of market risk is seen to switch from a low to a high regime, even though when it takes on the low value the probability of rejecting the null $H_0 : \lambda_{M1} = 0$ is equal to 90.78%. The price of intertemporal risk also show evidence of two states, both significantly different from zero. Since the Markov transition probabilities p and q are significant as well, one can appeal to the economic significance of a second regime, interpreted as a change in investors' preferences[25]. Interestingly, all the ζ's, apart from ζ_2, are significant, showing, once more, that the model proposed to capture asymmetry in second moments, though simplified, is still valid. Moreover, a Wald statistics which tests the null $H_0 : \zeta_1 = \ldots = \zeta_4 = 0$ rejects it at any conventional level. The Wald statistics is equal to 52.686 and it is

[25]Even though the model with switching prices of risk nests that with constant prices as a special case, to test for the statistical significance of the second regime, one cannot use a Likelihood Ratio (LR) test. This is due to the presence of nuisance parameters, i.e. those that are not identified under the null of a single regime. In this case the usual regularity conditions which justify the χ^2 approximation to the LR test no longer hold. Hansen (1992, 1996) has suggested asymptotically valid tests that overcome this difficulty. The procedure, however, is computationally burdensome and it is usually applied to quite simple models. Therefore it will not be adopted in this research.

asymptotically distributed as a χ^2_m, with $m = 4$ degrees of freedom. Its critical value at 99% confidence level is 13.277. Finally, all the k's turn out to be significant.

As for the switches in the prices of risk, two relevant issues can be discussed. The first is to uncover whether these changes are caused by movements in some macroeconomic fundamentals, like business cycle phases, or if they are generated by other factors, like changes in the states of financial market volatilities, which are not necessarily matched by contractions or expansions. The second is to check whether investors tend to become more risk averse before a financial turmoil, anticipating a future surge in volatility. To address these questions the *smoothed* probabilities become of interest. In estimating regime switching models, apart from the Markov transition probabilities, three conditional probabilities are usually computed: The *ex ante* probability, $P(S_{t+1} = k \,|\Im_t; \boldsymbol{\theta})$, $k = 1, \ldots, K$, which is of interest in forecasting and is based on an evolving information set; the filter probability, $P(S_t = k \,|\Im_t; \boldsymbol{\theta})$, which serves to filter the data and infer the likelihood of the current regime; finally, the smoothed probability, $P(S_t = k \,|\Im_T; \boldsymbol{\theta})$, which is based on the entire information set available and is designed to determine if and when the regime switches have occurred. In the top panel of Figure 3 the smoothed probabilities of the state in which the price of market risk is high while the price of intertemporal risk is low (state 2) are plotted. These are calculated with the algorithm developed by Kim (1994). In the bottom panel squared market excess returns (SMER) are represented. The latter is a rough indicator of the unconditional market volatility (Merton, 1980) and is computed as follows:

$$SMER_{t+1} = \left(\mathbf{w}'_{t+1}\mathbf{R}_{t+1}\right)^2,$$

where $\mathbf{w}_{t+1}$ is the vector of value weights and $\mathbf{R}_{t+1}$ the vector of asset excess returns. Vertical lines indicate the NBER business cycle peak and trough dates. Contrary to previous research (see, for instance, Harvey, 1989, Bekaert and Harvey, 1995, Chauvet and Potter, 1998), here it is not possible to detect a clear correspondence between business cycle contractions, on one hand, and increases in the coefficient of risk aversion, on the other hand. The probability of a high price of market risk, instead, seems to match periods of high market volatility, though, also in this case, the correlation is not that strong. Without considering the estimates obtained during earlier periods of the sample which are imprecise since little information from the data is used, the following regularities can be observed: The major spikes which hit financial markets in October 1974 and October 1987 are captured by the model but with some lags. The turbulence that occurred between 1979 and 1982, when the Fed ceased to target interest rates, is also not fully reflected by the

Table 6
Estimation results for the Merton's (1973) Intertemporal CAPM with time-varying prices of risk
Conditional second moment specification: Generalized GJR model

The estimated model is:

$$\mathbf{R}_{t+1} = \lambda_{M,S_t} \sum_{j=1}^{3} \mathbf{h}_{j,t+1} w_{j,t} + \lambda_{F,S_t} \mathbf{h}_{4,t+1} + \boldsymbol{\varepsilon}_{M,t+1},$$

$$IP_{t+1} = k_0 + k_1 IP_t + k_2 3MTR_{t-5} + \varepsilon_{IP,t+1},$$

The error terms are conditionally Student t-distributed, i.e. $\boldsymbol{\varepsilon}_{t+1} | \Im_t \sim Student - t \ (0, \mathbf{H}_{t+1}, \nu)$. The conditional covariance matrix follows the generalized GJR process described by equation (35):

$$\mathbf{H}_{t+1} = \mathbf{H}_0 \odot (\mathbf{11}' - \mathbf{aa}' - \mathbf{bb}') + \mathbf{aa}' \odot \overline{\boldsymbol{\varepsilon}}_t \overline{\boldsymbol{\varepsilon}}_t' + \boldsymbol{\zeta\zeta}' \odot \overline{\boldsymbol{\eta}}_t \overline{\boldsymbol{\eta}}_t' + \mathbf{bb}' \odot \mathbf{H}_t.$$

where $\overline{\boldsymbol{\varepsilon}}_t = \left[\overset{\wedge}{\boldsymbol{\varepsilon}}_{M,t} \ \varepsilon_{IP,t}\right]'$, $\overline{\eta}_{stock,t} = -I\left(\overset{\wedge}{\varepsilon}_{stock,t} < 0\right) \ \overset{\wedge}{\varepsilon}_{stock,t}$, $\overline{\eta}_{bill,t} = I\left(\overset{\wedge}{\varepsilon}_{bill,t} > 0\right) \ \overset{\wedge}{\varepsilon}_{bill,t}$, $\overline{\eta}_{bond,t} = \left[2 \cdot I\left(\overset{\wedge}{\varepsilon}_{bond,t} > 0\right) - 1\right] \ \overset{\wedge}{\varepsilon}_{bond,t}$, and $\eta_{IP,t} = [2 \cdot I(\varepsilon_{IP,t} > 0) - 1]\, \varepsilon_{IP,t}$.

λ_{M1}	λ_{M2}	λ_{F1}	λ_{F2}	p	q
0.022	0.082	−0.545	−3.978	0.963	0.942
(0.013)	(0.023)	(0.273)	(0.478)	(0.017)	(0.026)

k_0	k_1	k_2	ν
0.467	0.273	−0.484	4.463
(0.077)	(0.044)	(0.129)	(0.460)

	R^{stock}	R^{bill}	R^{bond}	IP
a_i	0.248	0.340	0.313	0.113
	(0.043)	(0.027)	(0.030)	(0.013)
ζ_i	0.354	−0.045	0.242	0.050
	(0.070)	(0.070)	(0.043)	(0.015)
b_i	0.916	0.936	0.945	0.994
	(0.019)	(0.011)	(0.010)	(0.003)

Log-likelihood function	– 7.300

Standard errors, based on the matrix of second derivatives of the log likelihood function, are shown in parentheses.

smoothed probabilities, which increased only in the 1982 spring/summer. This might be due to the composition of the portfolio analyzed here. Since, over the whole sample, the market portfolio never has less than 60% equity, the weight assigned to the conditional stock volatility dominates the covariances between stocks and the other two assets. Therefore, smoothed probabilities are mainly affected by stock market turmoils. An alternative interpretation is that investors become more risk averse after financial crises have unfolded. Apart from these events, any time squared market excess returns go up the probability of high risk aversion also increases. Such a result is in line with Campbell and Cochrane (1999), where risk aversion rises when asset prices fall. The fact that a high price of market risk is matched by a low price of intertemporal risk has an appealing intuition: The higher the risk aversion, the lower the fraction of wealth invested in risky assets, and, therefore, the less the importance attached to the intertemporal risk. Furthermore, even for the periods for which it is possible to establish a simultaneous link between an increase in financial market volatility and in the smoothed probabilities, it seems that these latter do not anticipate financial turmoil.

Thanks to the smoothed probabilities it is also possible to compute the relative likelihood of each regime with the formula proposed by Dahlquist and Gray (2000):

$$\overset{\wedge}{S_t} = \underset{k \in \{1,2\}}{\arg\max} \; P\left(S_t = k \,|\Im_T; \boldsymbol{\theta}\right). \tag{42}$$

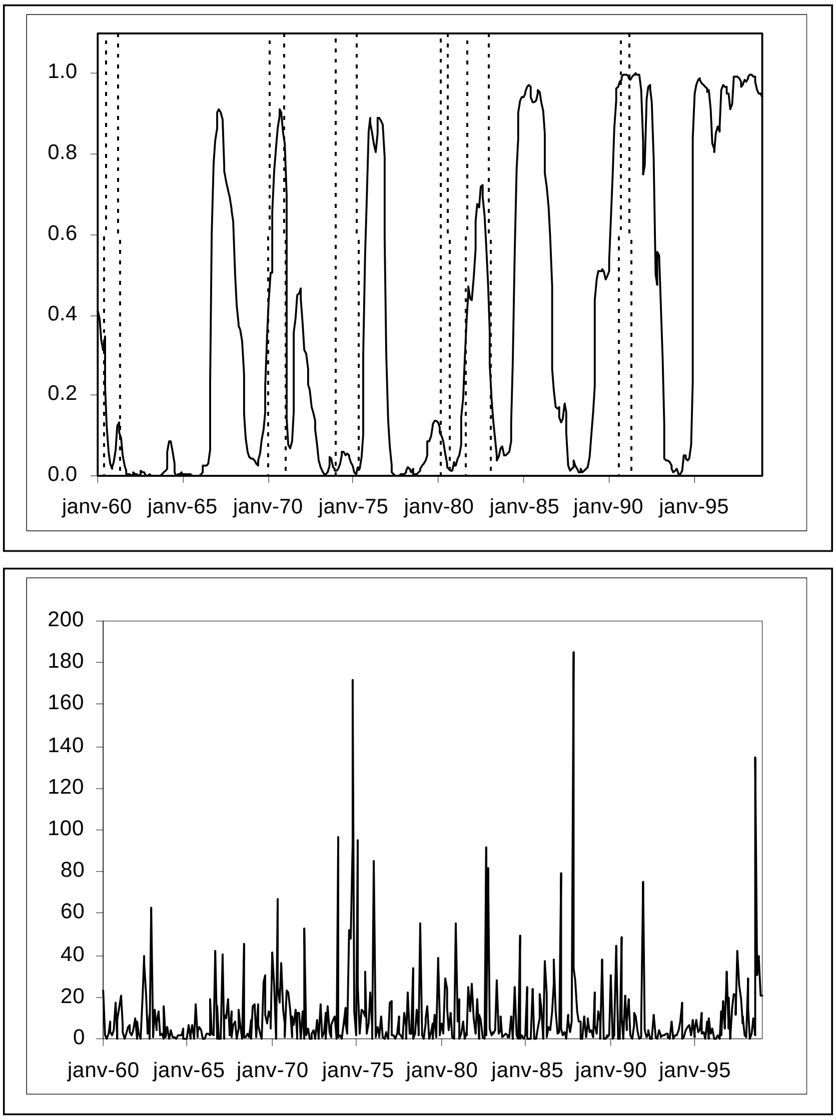
Equation (42), which assigns observations to a particular regime on the grounds of the relative probabilities of being in each regime, also allows one to estimate the number of switches. The observations which fall in the low risk aversion-high price of intertemporal risk state is equal to 300 (i.e. investors are in state 1 about 64% of the time), while there are 17 switches from one regime to the other. As reflected by the Markov probabilities, each regime is quite persistent. Conditional on being in state 1, the expected duration (in months) of low risk aversion-high price of intertemporal risk is (see Hamilton, 1989):

$$D_1 = \sum_{j=1}^{\infty} j q^{j-1} = \frac{1}{1-q} = 17.241.$$

Similarly, the expected duration of high risk aversion-low price of intertemporal risk, conditional on state 2, is

$$D_2 = \sum_{j=1}^{\infty} j p^{j-1} = \frac{1}{1-p} = 27.027.$$

Figure 3: Smoothed probabilities and squared market excess returns



Vertical lines indicate the NBER business cycle peak and trough dates.

While for the Intertemporal CAPM with constant prices of risk all the risk premia can be computed in a deterministic fashion, when a regime switching model is introduced, the premia have a stochastic component, given by the switching prices. Expected market, intertemporal, and total premia (EMRP, EIRP, and ETRP, respectively) for each asset are plotted in Figure 4 according to the following formulae:

$$EMRP_{i,t+1} = \left[\sum_{k=1}^{2} \lambda_{M,S_t=k} P\left(S_t = k \left| \Im_{t-1}; \boldsymbol{\theta} \right.\right) \right] \left[\sum_{j=1}^{3} Cov\left(R_{i,t+1}, R_{j,t+1} \left| \Im_t \right.\right) w_{j,t} \right], \tag{43}$$

$$EIRP_{i,t+1} = \left[\sum_{k=1}^{2} \lambda_{F,S_t=k} P\left(S_t = k \left| \Im_{t-1}; \boldsymbol{\theta} \right.\right) \right] Cov\left(R_{i,t+1}, IP_{t+1} \left| \Im_t \right.\right), \tag{44}$$

$$ETRP_{i,t+1} = EMRP_{i,t+1} + EIRP_{i,t+1}, \tag{45}$$

for $i = stock, bill, bond$. $P\left(S_t = k \left| \Im_{t-1}; \boldsymbol{\theta} \right.\right)$ represent the *ex ante* probabilities discussed in Section 4. For each asset, the time evolution of the three premia are quite similar to the path seen when prices of risk were held constant. It is the magnitude, rather the pattern of the premia, which varies, due to the change in the proportionality factors (i.e. the prices of risk). Therefore, the reader is referred to Section 5.1.2 for the interpretation of the results. Here, only the main differences are highlighted.

As far as stocks are concerned, market premia show much less pronounced spikes than in the case of constant prices of risk. The only exception occurs at the end of 1998, where the market premium increases much more when Hamilton's (1988, 1989, 1990, 1994) filter is employed. As expected, the differences in the intertemporal risk premium are more relevant due to the constraint $\zeta_{IP1} = \zeta_{IP2} = 1$ and go beyond a pure scale effect. Here, a huge trough occurs in the first half of the '70s, mainly due to a switch in the price of intertemporal risk, which moves from a high to a low state. While in the middle of the '70s the premium becomes positive when prices of risk are constant, it is negative when they are allowed to vary over time, a pattern which can be attributed to the different GARCH parametrizations. The remaining differences are more modest and can be essentially explained by the use of time varying prices of risk. Finally, notice that, also in this case, the market premium is the main component of the total premium.

The T-bill market and intertemporal risk premia follow approximately the same time path both when prices are kept constant and when they are allowed to vary. The scale effect seems to play the major role in explaining the differences between the two models.

As for the bond market risk premium, the most striking differences appear after the stock market crash in October 1987 and during the 1997/98 Asian-Russian-Latin American crises. The model where the constraint $\zeta_{bond1} = \zeta_{bond2} = 1$ is imposed is not able to capture the "flight to quality" from stocks to bonds which occurred at those times. Whereas the intertemporal risk premium plunges at the beginning of the '70s in the specification with time varying prices of risk, it is positive in the case where the constraint $\zeta_{IP1} = \zeta_{IP2} = 1$ is relaxed.

All in all, if the specification where prices of risk can change through time allows to accommodate shifts in investors' preferences, it sacrifices flexibility in terms of the GARCH parametrization. Relevant phenomena, like the decrease in the bond risk premia which occurred after some equity market falls, are not captured.

Due to the less satisfactory GARCH model as far as asymmetry is concerned, news impact curves and surfaces are not plotted here and the robust conditional moment test statistics of Kroner and Ng (1998) are not computed. The usual Ljung-Box portmanteau statistics for the standardized residuals in level, i.e. $\overset{\wedge}{\varepsilon}{}^{*}_{i,t+1} = \overset{\wedge}{\varepsilon}_{i,t+1}\ h_{ii,t+1}^{-1/2}$, and squared, i.e. $\overset{\wedge}{\varepsilon}{}^{*2}_{i,t+1} = \overset{\wedge}{\varepsilon}{}^{2}_{i,t+1}\ h_{ii,t+1}^{-1}$, are, instead, proposed and reported in Table 6[26]. As in the case where prices of risk are held constant, no simultaneous autocorrelation for stocks and bonds is detected for residuals either in level or squared. The null hypothesis is, instead, rejected for T-bill error terms, suggesting that an autoregressive term in the T-bill mean equation might provide a better fit. Extra explanatory variables are probably called for with regards to the second moment specification of the industrial production growth rate as well, since the Ljung-Box portmanteau statistics for squared residuals is significant at 1% level. Notice that, when the constraint $\zeta_{IP1} = \zeta_{IP2} = 1$ is relaxed, IP squared residuals do not show any autocorrelation (see Section 5.1.3).

The model with time-varying prices of risk has also been extended by allowing the Markov switching probabilities to be state-dependent. In particular equation (23) has been specified as follows:

$$\begin{aligned} p_t &= \Phi\left(\beta_{p0} + \beta_{p1} Z_{t-1}\right) \\ q_t &= \Phi\left(\beta_{q0} + \beta_{q1} Z_{t-1}\right) \end{aligned}, \tag{46}$$

where $\Phi(\cdot)$ is the cumulative normal distribution and Z_{t-1} is an information

[26] Notice that when $i = IP$, residuals do not need to be averaged out (See Section 4.2 for further details).

Figure 4: Components of excess returns

Conditional Intertemporal CAPM with time-varying prices of risk. Conditional second moment specification: Generalized GJR model. The scale along the vertical axes is in units of percent.

Figure 4A: Stock excess returns

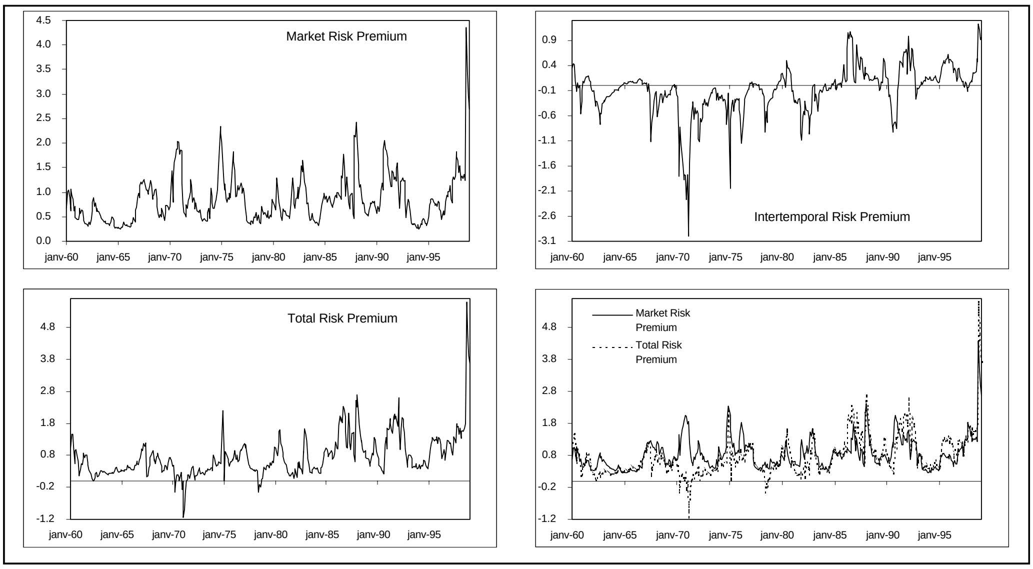


Figure 4B: Excess holding yields on 6-month T-bills

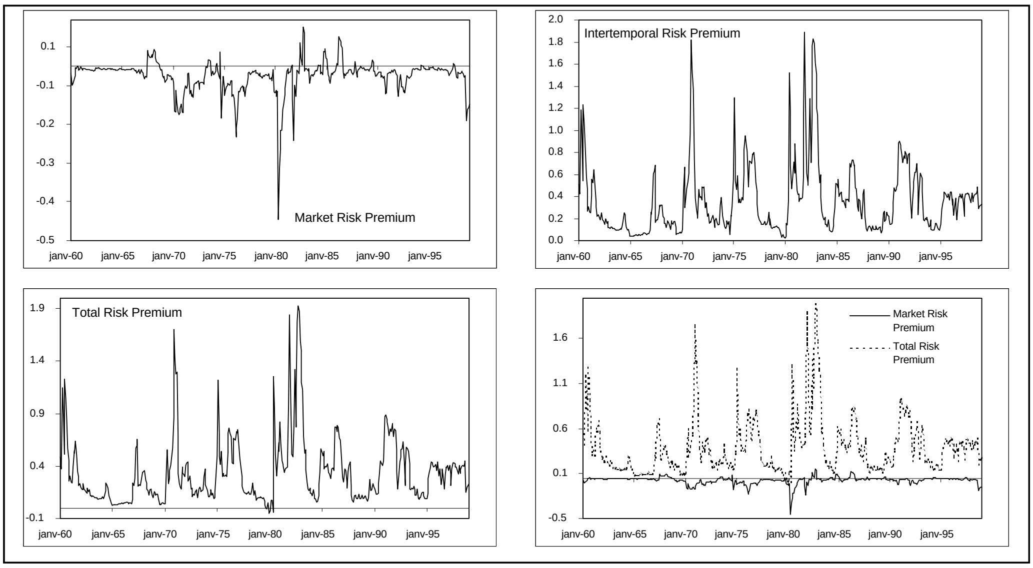


Figure 4C: Excess holding yields on 10-year government bonds

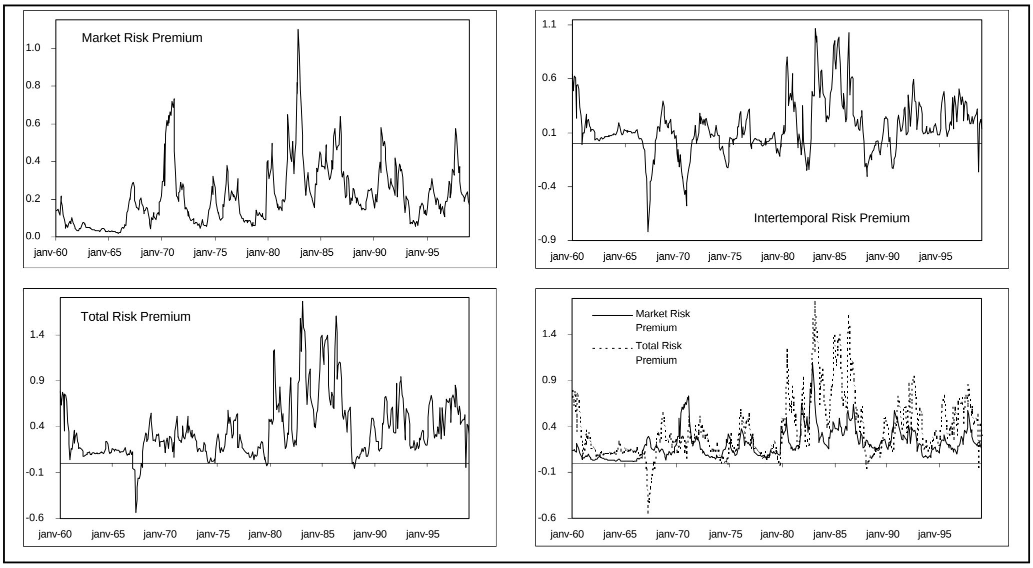


Table 6
Diagnostic tests for the standardized residuals of the generalized GJR model

	R^{stock}	R^{bill}	R^{bond}	IP
Mean	-0.051	0.080	-0.097	-0.023
Std. Dev.	0.846	0.979	0.822	0.870
$\mathbf{L\text{-}B_{12}}\left(\overset{\wedge}{\varepsilon}{}^{*}_{i,t+1}\right)$	7.573	28.029^{**}	8.020	15.500
$\mathbf{L\text{-}B_{12}}\left(\overset{\wedge}{\varepsilon}{}^{*2}_{i,t+1}\right)$	6.232	39.066^{**}	13.420	34.878^{**}

** denotes 1% significance level.
The Ljung-Box$_m$ (L-B$_m$) statistics tests the null hypothesis that all autocorrelation coefficients are simultaneously equal to zero up to m lags; it is asymptotically distributed as χ^2_m. It has been chosen $m = 12$ for which the critical values at 95% and 99% confidence level are 21.026 and 26.217, respectively.
Notice that when $i = IP$, residuals do not need to be averaged out (see Section 4.2 for further details).

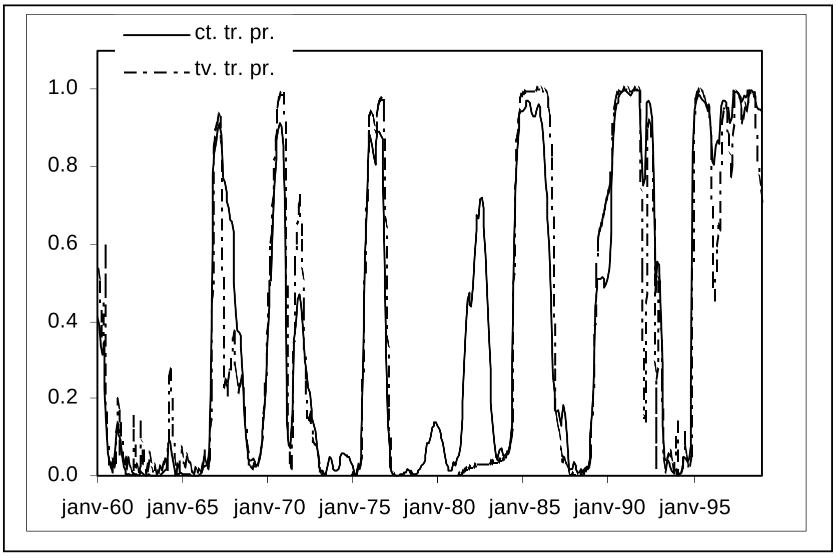
variable. A large set of instruments has been experimented[27]. The one that has provided with the most sensible results is the one-year Treasury bill yield, $1YR_{t-1}$ (source: Federal Reserve). The choice of $1YR_{t-1}$ is motivated by the fact that the switch to the high-risk aversion/low-price of intertemporal risk regime may be more likely when, for instance, that interest rate becomes high. In these times, typically, stock markets tend to fall, industrial production usually decreases and agents become more risk-averse. However, since a Likelihood Ratio (LR) test clearly rejects the model with state-dependent transition probabilities, parameter estimates for this specification are not reported. The LR statistics is equal to 0.020 and is asymptotically distributed as a χ^2_m, with $m = 2$ degrees of freedom. Its critical value at the 95% confidence level is 5.991. Nevertheless, the smoothed probabilities relative to the regime where the coefficient of risk aversion is high and the price of intertemporal risk is low are reported and compared with the smoothed probabilities of the same state obtained imposing constant Markov transition probabilities. As evident from Figure 5, the more important difference between the time evolution of the two smoothed probabilities occurs in the first half of the '80s: When the Markov switching probabilities are allowed to vary over time, the correspondent smoothed probabilities are quite flat. This model, thus, fails to capture the turbulence in the government fixed income security market due to the change in the Fed's monetary policy.

6 Summary of Results and Conclusions

This paper estimates a trivariate two-factor conditional version of the Intertemporal CAPM of Merton (1973). The analysis considers three assets: US stocks, 6-month T-bills, and 10-year government bonds, with the purpose of documenting asymmetry in conditional second moments also for fixed income securities. As a second factor the growth rate of industrial production is chosen. The approach followed here allows to distinguish between market and intertemporal premia for each asset. The second factor turns out to

[27]The information variables used are: 1) The dividend yield on the S&P 500 composite equity index (source: S&P - DRI). 2) The stock market returns computed on the S&P composite equity index (source: S&P - DRI). 3) The change in the default spread, defined as Moody's Baa minus Aaa long-term bond rate (source: Federal Reserve). 4) The term premium, defined as the return on 10-year government bonds minus the 3-month T-bill rate (source: CRSP). 5) The default premium, defined as the return on a portfolio of low-grade corporate bonds less the return on a portfolio of long-term government bonds (source for the low-grade corporate bond series: Ibbotson Corporate Bond Module portfolio of bonds rated Baa or less by Moody's Investor Services; Blume, Keim, and Patel (1991); and Salomon Brothers. Source for long-term government bonds: Salomon Brothers total return Treasury index). 6) and 7) The return on the 1-/3-month T-bill (source: CRSP). 8) The 1-year Treasury bill yield (source: Federal Reserve).

Figure 5: Smoothed probabilities with and without time-varying Markov transition probabilities



"ct. tr. pr." stands for constant transition probabilities; "tv. tr. pr." stands for time-varying transition probabilities.

be significantly priced and relevant in the determination of the total premia required to hold risky securities.

Engle and Ng (1993) tests for asymmetry show that the volatility of the three assets as well as that of the second priced factor are sensitive to the sign of past innovations. Moreover, the volatility of each variable responds differently to past positive and/or negative shocks. Therefore two multivariate GARCH processes able to capture the asymmetric effects for both conditional variances *and* covariances are developed and tested. While the generalized Nelson model is not robust to different sets of starting values and numerical algorithms of optimization, the generalized GJR model seems to fit the data well.

Major financial market turmoil as well as spillovers from one market to another are reflected by the time path of risk premia. More importantly, as evidenced by plots of the news impact curves and surfaces and confirmed by robust conditional moment tests (Engle and Ng, 1993, and Kroner and Ng, 1998), conditional second moments do respond asymmetrically to past positive and/or negative news. Traditional GARCH models do not take into account such asymmetric effects. Therefore, poor in- and out-of-sample forecasts as well as misleading indications for portfolio selection, risk management, and pricing of primary and derivative securities can be the consequence of this omission.

Following Cappiello (1999), the prices of risk are first held constant and next allowed to vary over time according to the regime switching model of Hamilton (1988, 1989, 1990, 1994). Due to its better performance, the use of Hamilton's filter is combined only with the generalized GJR model. Two regimes are identified, one in which the price of market risk is high and the price of intertemporal risk is low, and one in which the reverse occurs, i.e. the first price is high and the second low. This can be interpreted as a switch in investors' preferences whose degree of risk aversion increases in correspondence to or after financial turmoil. The shift in agents' preferences, though, is accommodated at the cost of sacrificing flexibility in terms of the GARCH specification. Thus, relevant phenomena, like the decrease in the bond risk premia, which occurred after some equity market falls, are not captured, whereas they are when prices are held constant.

Several directions for future research can be pursued.

Due to the partial equilibrium nature of Merton's (1973) Intertemporal CAPM, what pricing factors should be considered remain an open question and further investigation is called for. Therefore, the results obtained here choosing the growth rate of industrial production as second priced factor, must be interpreted with caution.

In general, GARCH models, both symmetric and asymmetric, are statistical specifications designed to capture some empirical regularities typical of financial time series. However, they lack economic content. A satisfactory dynamic theoretical model able to explain the asymmetric responses of a large spectrum of assets to past innovations is definitely called for. The theories proposed by Black (1976) and Christie (1982), on one hand, and Campbell and Hentschell (1992), on the other, though appropriate for equities, do not provide exhaustive explanations for (government) fixed income securities.

Between the two multivariate GARCH processes developed in this research, the generalized Nelson model does not seem to be suited for fixed income securities. Therefore, it can be tested again with portfolios made of equities only. This would allow to check whether, contrary to the results reached in this paper, it is robust to different sets of starting values and numerical algorithms of optimization and if it is stationary.

In the generalized GJR model the component which captures asymmetry in second moments shows an abrupt transition from one volatility regime to another. Instead of using the indicator function, a smooth transition function could be employed in the spirit of Lubrano (1998), where the smooth transition proved to be relevant in the empirical analysis.

As shown by Engle and Ng (1993), when portfolios of only equities are considered, after a price drop, which triggers an increase in volatility, traditional GARCH models underestimate the conditional variance. Similarly, after a stock price rise, symmetric GARCH overestimate the conditional volatility. Hence the poor forecasting performance of symmetric GARCH processes and the need of asymmetric GARCH specifications. The problem of poor in- and out-of-sample forecasting of traditional GARCH processes has also been addressed in an alternative way. The estimation of conventional GARCH models typically show very high persistence parameters (close to one). Such a persistence has been judged spurious and caused by structural shifts in the level of the conditional covariance not accounted for by the econometrician (see Diebold, 1986, and Lamoureux and Lastrapes, 1990). The fact that traditional GARCH models overpredict the conditional volatility at the time when it goes from a high state back to a low (or normal) state, and similarly underpredict it when volatility changes from low to high (or normal) states has been attributed to this high spurious persistence. Among others, Cai (1994), Hamilton and Susmel (1994), Dueker (1997), and Susmel (1998) capture the shifts in the conditional variance through the regime switching model of Hamilton (1988, 1989, 1990, 1994). When such shifts are taken into account, the persistence decreases and the forecasting performance improves. An interesting future project would compare the two approaches, i.e. multivariate asymmetric *versus* switching GARCH models and test which offers the best results in terms of forecasting.

Finally, note that Merton's (1973) Intertemporal CAPM can be used as a tool for asset allocation. The maximization process of any CAPM produces optimal weights for the allocation of wealth among different asset classes. The weights are a function of the lagged covariance matrix, prices of risk, and expected excess returns. These three terms can be estimated through Merton's asset pricing theory. In addition, since a conditional version of the model is considered, both first and second moments vary over time. As a consequence, investors' portfolios can be rebalanced as new information becomes available.

References

Baillie, R. T., and R. P. DeGennaro, 1990, "Stock Returns and Volatility," *Journal of Financial and Quantitative Analysis*, 25, pp. 203-214.

Bekaert, G., and C. R. Harvey, 1995, "Time-Varying World Market Integration," *Journal of Finance,* 50, pp. 403-444.

Bekaert, G., and C. R. Harvey, 1997, "Emerging Equity Market Volatility," *Journal of Financial Economics,* 43, pp. 29-77.

Bekaert, G., and G. Wu, 2000, "Asymmetric Volatility and Risk in Equity Markets," *Review of Financial Studies*, 13, pp. 1-42.

Bera, A. K., and M. L. Higgins, 1993, "ARCH Models: Properties, Estimation, and Testing," *Journal of Economic Surveys*, 7, pp. 305-366.

Berndt, E. K., B. H. Hall, R. E. Hall, and J. A. Hausman, 1974, "Estimation and Inference in Non-Linear Structural Models," *Annals of Economic and Social Measurement* 3, pp. 653-665.

Black, F., 1976, "Studies of Stock Price Volatility Changes," *Proceedings from the American Statistical Association, Business and Economic Statistics Section*, pp. 177-181.

Bollerslev, T., 1986, "Generalized Autorregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, pp. 307-327.

Bollerslev, T., 1987, "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, 69, pp. 542-547.

Bollerslev, T., 1990, "Modeling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach," *Review of Economics and Statistics*, 72, pp. 498-505.

Bollerslev, T., R. Y. Chou, and K. F. Kroner, 1992, "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, pp. 5-59.

Bollerslev, T., R. F. Engle, and D. B. Nelson, 1994, "ARCH Models," in: R. F. Engle, and D. L. McFadden (eds.), *Handbook of Econometrics*, vol.4, Elsevier, pp. 2959-3038.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge, 1988, "A Capital Asset Pricing Model with Time Varying Covariances," *Journal of Political Economy*, 96, pp. 116-131.

Bonomo, M., and R. Garcia, 1997, "Tests of Conditional Asset Pricing Models in the Brazilian Stock Market," *CIRANO working paper*, # 97s-20.

Booth, G. G., and G. Koutmos, 1995, "Asymmetric Volatility Transmission in International Stock Markets," *Journal of Banking and Finance*, 21, pp. 811-823.

Booth, G. G., T. Martikainen, and Y. Tse, 1997, "Price and Volatility Spillovers in Scandinavian Stock Markets," *Journal of International Money and Finance*, 14, pp. 747-762.

Braun, P. A., D. B. Nelson, and A. M. Sunier, 1995, "Good News, Bad News, Volatility, and Betas," *Journal of Finance*, 50, pp. 1575-1603.

Cai, J., 1994, "A Markov Model of Unconditional Variance in ARCH," *Journal of Business and Economic Statistics,* 12, pp. 309-316.

Campbell, J. Y., 1987, "Stock Returns and the Term Structure," *Journal of Financial Economics*, 18, pp. 373-399.

Campbell, J. Y., and J. H. Cochrane, 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, pp. 205-251.

Campbell, J. Y., and L. Hentschell, 1992, "No News Is Good News," *Journal of Financial Economics*, 31, pp. 281-318.

Cappiello, L., 1998, "Measuring Exchange Rate Risk Premia: the Case of the USA and Italy," *working paper*, International Economics Section, The Graduate Institute of International Studies, University of Geneva.

Cappiello, L., 1999, "A Conditional Multi-asset Intertemporal CAPM with Switching Prices of Risk," *paper included in the Ph.D. Dissertation*, International Economics Section, The Graduate Institute of International Studies, University of Geneva.

Chauvet, M., 1999, "Stock Market Fluctuations and the Business Cycle," *working paper*, Department of Economics, University of California, Riverside.

Chauvet, M., and S. Potter, 1998, "Non-Linear Risk," *working paper*, Department of Economics, University of California, Riverside.

Chen, N.-F., R. Roll, and S. A. Ross, 1986, "Economic Forces and Stock Market," *Journal of Business*, 59, pp. 383-403.

Cheung, Y-W., and L. K. Ng, 1992, "Stock Price Dynamics and Firm Size: An Empirical Investigation," *Journal of Finance*, 47, pp. 1985-1997.

Chou, R., R. F. Engle, and A. Kane, 1992, "Measuring Risk Aversion from Excess Returns on a Stock Index," *Journal of Econometrics*, 52, pp. 201-224.

Christie, A. A., 1982, "The Stochastic Behaviour of Common Stock Variances: Value, Leverage, and Interest Rate Effects," *Journal of Financial Economics*, 10, pp. 407-432.

Cochrane, J. H., 1999, "New Facts in Finance," *NBER working paper*, # 7169.

D'Agostino, R. B., A. Belanger, and R. B. D'Agostino Jr., 1990, "A Suggestion for Using Powerful and Informative Tests of Normality," *The American Statistician,* 44, pp. 316-321.

Dahlquist, M., and S. F. Gray, 2000, "Regime-Switching and Interest Rates in the European Monetary System," *Journal of International Economics,* forthcoming.

De Santis, G., and B. Gerard, 1997, "International Asset Pricing and Portfolio Diversification with Time-Varying Risk," *Journal of Finance,* 52, pp. 1881-1912.

De Santis, G., and B. Gerard, 1998a, "How Big is the Premium for Currency Risk?," *Journal of Financial Economics* 49, pp. 375-412.

De Santis, G., and B. Gerard, 1998b, "Conditional Asset Pricing with Market and Inflation Risk," *working paper*, Marshall School of Business, University of Southern California.

De Santis, G., B. Gerard, and P. Hillion, 1998, "The Relevance of Currency Risk in the EMU," *working paper*, Marshall School of Business, University of Southern California.

De Santis, G., and S. İmrohoğlu, 1997, "Stock Returns and Volatility in Emerging Financial Markets," *Journal of International Money and Finance*, 16, pp. 561-579.

Diebold, F. X., 1986, "Comment on Modelling the Persistence of Conditional Variance," *Econometric Reviews,* 5, pp. 51-56.

Diebold, F. X., J. H. Lee, and G. C. Weinbach, 1994, "Regime Switching with Time-Varying Transition Probabilities," in: C. Hargreaves (ed.), *Nonstationary Time Series Analysis and Cointegration*, Oxford University Press, pp. 283-302.

Ding, Z., and R. F. Engle, 1994, "Large Scale Conditional Covariance Matrix Modelling, Estimation and Testing," *working paper*, University of California, San Diego.

Ding, Z., C. Granger, and R. F. Engle, 1993, "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance*, 1, pp. 83-106.

Dumas, B., and B. Solnik, 1995, "The World Price of Foreign Exchange Risk," *Journal of Finance,* 50, pp. 445-479.

Dueker, M. J., 1997, "Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility," *Journal of Business and Economic Statistics,* 15, pp. 26-34.

Duffee, G. R., 1995, "Stock Returns and Volatility: A Firm Level Analysis," *Journal of Financial Economics*, 37, pp. 399-420.

Durland, J. M., and T. H. McCurdy, 1994, "Duration Dependent Transitions in a Market Model of US GNP Growth," *Journal of Business and Economic Statistics*, 12, pp. 279-288.

Engle, R. F., 1990, "Discussion: Stock Volatility and the Crash of '87," *Review of Financial Studies*, 3, pp. 103-106.

Engle, R. F., and K. F. Kroner, 1995, "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11, pp. 122-150.

Engle, R. F., and V. K. Ng, 1993, "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48, pp. 1749-1778.

Engle, R. F., V. K. Ng, and M. Rothschild, 1990, "Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills," *Journal of Econometrics*, 45, pp. 213-238.

Fama, E. F., 1965, "The Behaviour of Stock Prices," *Journal of Business*, 38, pp. 34-105.

Fama, E. F., 1998, "Determining the Number of Priced State Variables in the ICAPM," *Journal of Financial and Quantitative Analysis*, 33, pp. 217-231.

Fang, K.-T., S. Kotz, and K. W. Ng, 1990, *Symmetric Multivariate and Related Distributions*, Chapman and Hall.

Filardo, A.J., 1994, "Business Cycle Phases and their Transitional Dynamics," *Journal of Business and Economic Statistics*, 12, pp. 299-308.

Fornari, F., and A. Mele, 1997, "Sign- and Volatility Switching ARCH Models: Theory and Applications to International Stock Markets," *Journal of Applied Econometrics,* 12, pp. 49-65.

French, K. R., W. Schwert, and R. F. Stambaugh, 1987, "Expected Stock Returns and Variance," *Journal of Financial Economics*, 19, pp. 3-29.

Friedman, B. M., and K. N. Kuttner, 1992, "Time-Varying Risk Perceptions and the Pricing of Risky Assets," *Oxford Economic Papers*, 44, pp. 566-598.

Ghysels, E., 1993, "On Periodic Stochastic Regime Switching in Time Series Models," *Journal of Business and Economic Statistics,* 12, pp. 289-298.

Glosten, L. R., R. Jagannathan, 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, pp. 1779-1801.

Gouriéroux, C., and A. Monfort, 1992, "Qualitative Threshold ARCH Models," *Journal of Econometrics*, 52, pp. 159-199.

Gouriéroux, C. and A. Monfort, 1995, *Statistics and Econometric Models*, Cambridge University Press.

Gray, S. F., 1996, "Modelling the Conditional Distribution of Interest Rates as a Regime Switching Process," *Journal of Financial Economics,* 42, pp. 27-62.

Hamilton, J. D., 1988, "Rational-Expectation Econometric Analysis of Changes in Regimes - An Investigation of the Term Structure of Interest Rates", *Journal of Economic Dynamics and Control*, 12, pp. 385-423.

Hamilton, J. D., 1989, "A New Approach to the Economic Analysis of Non-Stationary Time Series and the Business Cycle," *Econometrica,* 57, pp. 357-384.

Hamilton, J. D., 1990, "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics,* 4, pp. 39-70.

Hamilton, J. D., 1994, *Time Series Analysis*, Princeton University Press.

Hamilton, J. D., and G. Lin, 1996, "Stock Market Volatility and the Business Cycle," *Journal of Applied Econometrics,* 11, pp. 573-593.

Hamilton, J. D., and R. Susmel, 1994, "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics,* 64, pp. 307-333.

Hansen, B. E., 1992, "The Likelihood Ratio Test under Non-Standard Conditions: Testing the Markov Trend Model of GNP," *Journal of Applied Econometrics,* 7, pp. S61-S82.

Hansen, B. E., 1996, "Erratum: The Likelihood Ratio Test under Non-Standard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics,* 11, pp. 195-198.

Harvey, C. R., 1989, "Time-Varying Conditional Covariances in Tests of Asset Pricing Models," *Journal of Financial Economics,* 24, pp. 289-317.

Harvey, C. R., 1991, "The World Price of Covariance Risk," *Journal of Finance,* 46, pp. 111-157.

Harvey, C. R., and A. Siddique, 1999, "Autoregressive Conditional Skewness," *Journal of Financial and Quantitative Analysis,* 34, pp. 465-488.

Hentschell, L., 1995, "All in the Family: Nesting Symmetric and Asymmetric GARCH Models," *Journal of Financial Economics*, 39, pp. 71-104.

Johnson, N. L., and S. Kotz, 1972, *Distributions in Statistics: Continuous Multivariate Distributions*, John Wiley & Sons.

Kim, C. J., 1994, "Dynamic Linear Models with Markov Switching," *Journal of Econometrics,* 64, pp. 1-22.

Koutmos, G., 1996, "Modeling the Dynamic Interdependence of Major European Stock Markets," *Journal of Business, Finance, and Accounting,* 23, pp. 975-988.

Koutmos, G., and G. G. Booth, 1995, "Asymmetric Volatility Transmission in International Stock Markets," *Journal of International Money and Finance,* 14, pp. 747-762.

Kroner, K. F., and V. K. Ng, 1998, "Modeling Asymmetric Comovements of Asset Returns," *Review of Financial Studies,* 11, pp. 817-844.

Lamoureux, C. G. and W. D. Lastrapes, 1990, "Persistence in Variance, Structural Change, and the GARCH Model," *Journal of Business & Economic Statistics,* 8, pp. 225-234.

Lubrano, M., 1998, "Smooth Transition GARCH Models: A Bayesian Perspective," *working paper*, GREQAM-CNRS, Marseille.

Malkiel, B. G., 1979, "The Capital Formation Problem in the United States," *Journal of Finance*, 40, pp. 677-687.

Mandelbrot, B., 1963, "The Variation of Certain Speculative Prices," *Journal of Business*, 36, pp. 394-419.

Merton, R. C., 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, pp. 867-887.

Merton, R. C., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics,* 8, pp. 323-361.

Nelson, D. B., 1991, "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, pp. 347-370.

Perez-Quiros, G., and A. Timmermann, 1999, "Firm Size and Cyclical Variation in Stock Returns," *working paper*, Department of Economics, University of California, San Diego.

Pindyck, R. S., 1984, "Risk, Inflation, and the Stock Market," *American Economic Review*, 74, pp. 335-351.

Poterba, T. M., and L. H. Summers, 1986, "The Persistence of Volatility and Stock Market Fluctuations," *American Economic Review,* 76, pp. 1142-1151.

Rabemananjara, R., and J. M. Zakoian, 1993, "Threshold ARCH Models and Asymmetries in Volatility," *Journal of Applied Econometrics,* 8, pp. 31-49.

Schwert, G. W., 1990, "Stock Volatility and the Crash of '87," *Review of Financial Studies*, 3, pp. 77-102.

Scruggs, J. T., 1998, "Resolving the Puzzling Intertemporal Relation between the Market Risk Premium and Conditional Market Variance: A Two-Factor Approach," *Journal of Finance*, 58, pp. 575-603.

Sentana, E., 1991, "Quadratic ARCH Models: A Potential Reinterpretation of ARCH Models as Second-Order Taylor Approximations," *working paper*, London School of Economics and Political Science, London.

Susmel, R., 1998, "Switching Volatilities in International Equity Markets," *working paper*, Department of Finance, University of Houston.

Turner, C. M., R. Startz, and C. R. Nelson, 1989, "A Markov Model of Heteroskedasticity Risk, and Learning in the Stock Market," *Journal of Financial Economics*, 25, pp. 3-22.

Wu, G., 2000, "The Determinants of Asymmetric Volatility," *Review of Financial Studies*, forthcoming.

Wu, G., and Z. Xiao, 2000, "A Generalized Partial Linear Model of Asymmetric Volatility," *working paper*, University of Michigan Business School, Ann Arbor.

Zakoian, J. M., 1994, "Threshold Heteroskedastic Models," *Journal of Economic Dynamics and Control*, 18, pp. 931-955.