Stochastic Volatility Risk and the Size Anomaly*

Claudia E. Moise
Weatherhead School of Management
Case Western Reserve University

April 12, 2007

Abstract

Investors' concerns about systematic volatility risk explain a large portion of the small firm premia in the long run. This novel finding supports the concept of market efficiency and indicates a "flight to quality" during recessions: investors shift their preferences away from small firms, which are considered as being relatively risky. Instead they use large, "quality" stocks, whose returns co-vary positively with innovations in volatility (a recession-and-distress proxy), and therefore pay off during times of low market returns. This leads to higher hedging demands for large stocks, higher prices and lower expected returns. Using a sample of monthly returns spanning the period January 1927-December 2005, I estimate a statistically significant and negative price for volatility risk. This result is robust to the type of volatility measure used, to the inclusion of traditional risk factors in the model and to different model specifications. Controlling for accounting profitability type, I document a significant volatility premium ranging between 3% and 7% per year, for the Fama-French portfolios.

JEL Classification: G12

Keywords: Volatility premium, Size anomaly, Cross-section of stock returns, Griddy-Gibbs sampler, GARCH models

*I gratefully acknowledge Kerry Back, Federico Bandi, John Campbell, John Cochrane, Eugene Fama, Michael Gallmeyer, Anurag Gupta, John Heaton, Hedibert Lopes, Nick Polson, Matt Pritsker, Myron Scholes, participants at the Finance Seminars at Case Western Reserve, George Mason, McGill, Texas A&M, Oak Hill Platinum Partners, the Asset Pricing Brownbag at the University of Chicago and the 4th Annual Trans-Atlantic Doctoral Conference at London Business School for their helpful comments. I thank Eugene Fama and Kenneth French for making the data on their factors and portfolios available. All errors are my own. Weatherhead School of Management, Case Western Reserve University, 10900 Euclid Avenue, 369 Peter B. Lewis Bldg., Cleveland, OH 44106. E-mail: claudia.moise@case.edu.

1 Introduction

It has been documented that stock market volatility is priced in the cross-section of stock returns. In particular, Ang et al. (2006) and Moise (2002) show that (innovations in) volatility is negatively priced in the cross-section of returns for stocks sorted based on past sensitivity to this factor. Is there additional information in volatility that is relevant for explaining documented return anomalies? In particular, can the size and value anomalies be explained by exposure to this risk factor, which has been postulated to be a recession-and-distress proxy [1]? This is the primary question I address in this paper.

Specifically, I test the following joint hypothesis. First, if volatility is a bad time proxy, as it has been postulated, then it should be priced in the cross-section of returns of the size- and value-sorted portfolios of Fama and French (1992, 1993) [2]. Second, if small- and value-firm premia are related to volatility risk, then the pattern in returns for the size- and value-sorted portfolios should match the pattern in their corresponding volatility loadings.

Consistent with my hypothesis and in agreement with existing literature, I find that returns are cross-sectionally related to their sensitivity to fluctuations in volatility. Using a sample of monthly returns spanning the period January 1927-December 2005, I find a statistically significant and negative price for volatility risk. More importantly, I document a pattern in returns for the size-sorted portfolios which matches the pattern in their corresponding volatility loadings. In addition, I show that growth firms have higher loadings on volatility than do value firms.

These results are appealing because they provide a risk-based interpretation for the size (and value) anomaly[3]. The returns of large stocks co-vary positively with innovations in volatility. Therefore, these stocks pay off during bad economic times, when volatility is more volatile and market returns are lower. Thus, they are perceived as a means of intertemporal hedge. According to rational pricing, this leads to higher hedging demands, higher prices and lower expected returns for large stocks. The differential volatility

[1] Seminal references include Banz (1981) and Reinganum (1981) for the size effect, Graham and Dodd (1934), Basu (1977, 1983), Ball (1978), and Rosenberg, Reid, and Lanstein (1985) for the value effect, and Fama and French (1992) for both effects.

[2] See Hamilton and Lin, 1996.

[3] Small and value firms consistently deliver higher returns. Since the spread in market betas between small and large, as well as the one between value and growth firms, is too small to account for the difference in their returns, these phenomena have been labeled as the size and value anomalies, respectively.

loadings between growth and value firms also lead to higher hedging demands and lower expected returns for growth stocks. Conversely, small and value stocks expose imvestors to greater volatility risk.

From an economic perspective, my results are consistent with a "flight to quality" explanation in a general equilibrium model. When faced with a downturn in the economy, investors are willing to forgo expected returns in order to get downside protection. Therefore, they shift their preferences towards low risk firms, which are large companies, often more established and with stronger relationships with their lenders, and growth companies, which are considered as being more profitable. The latter finding is in agreement with Petkova and Zhang (2005), who also document that value firms are riskier during bad times.

My hypothesis is motivated by two lines of research. The first one is related to volatility's pricing implications. Having a strong countercyclical pattern (it peaks just before or during recessions, and it falls sharply late in recessions or early in recovery periods), volatility has been considered a systematic source of risk (see Chen, 2003, Vayanos, 2004, Ang et al., 2006).

The other area of research leading to my hypothesis posits that there is a systematic difference between the variation in expected returns for small and large firms over economic cycles (see Perez-Quiros and Timmermann, 2000). Specifically, as recession approaches and business activity slows down, lower liquidity, higher short-term interest rates and higher default premia have a larger impact on small firms versus big firms, since the latter have more collateral. During the same times, firm profitability is more sensitive to adverse fluctuations in the market. Since it has been documented that value firms have persistently poor earnings (low stock price relative to book value, Fama and French, 1995), investors perceive both small and value firms as facing greater cyclical risk.

Therefore, I combine the two areas of research and test the joint hypothesis posited above. The sample period I use is characterized by several recessions and financial crises, which are very important for detecting downside risk. Analyzing volatility pricing implications over such a long period justifies its role as a state variable proxying for recessions and financial distress. As in the existing literature, I consider innovations in volatility as the factor of interest. The motivation for looking at innovations in volatility rather than its level comes from Merton (1973), who suggests that changes in variables that forecast future market returns should be used in explaining the cross-section of average returns, since they capture changes in the investment opportunities set.

The economic setting for this study is a pure exchange economy (Lucas,

1978) with no frictions, and a rational and infinitely lived representative consumer who has the recursive preferences proposed by Epstein and Zin (1989, 1991). As in Moise (2002), I invoke a 2-factor asset pricing model consisting of the return on the market portfolio and innovations in its volatility, which is theoretically based on Campbell's (1993) work and has its economic motivation rooted in Merton's ICAPM (1973).

My study contributes to the financial literature in several ways. First, I show that investors' concerns about systematic volatility risk explain a large part of the small firm premia in the long run (January 1927 - December 2005). This novel finding has immediate implications which support the concept of market efficiency, by establishing a link between volatility risk and premia associated with firm characteristics. I document that the observed size (and value) effect is not due to a misassessment of risk. It is driven by exposure to recession-and-distress risk (proxied here by innovations in volatility), which entails a significant premium. This premium ranges from 3% to 7% per year, across value-sorted portfolios. I find that, on average, over 90% of the size anomaly is explained by exposure to volatility risk.

Second, I provide evidence that volatility is a systematic source of risk which is negatively priced in the cross-section of returns, in the long run. This finding is robust with respect to the type of volatility measure used (ex-post or ex-ante) and to model specification. I build ex-ante volatility using an $Asymmetric\ Student-GARCH$ model which captures documented stylized facts about market volatility (i.e., clustering, autoregression, heteroskedasticity) and supports the leverage hypothesis of Black (1976). I provide additional evidence that my results are not driven by some missing risk factors by showing that the volatility effect is robust with respect to the size, value and momentum effects. I also show that my results hold under different model specifications.

Third, the heteroskedasticity and autocorrelation patterns from the time series of stock returns could lead to biased standard errors for the estimates of the factors' loadings in a simple OLS framework; therefore, I propose a novel specification for the time series asset pricing model ($IGARCH$), which results in more accurate statistical inferences.

I structure the rest of the paper as follows. In Section 2, I present the model and its economic motivation. In Section 3, I price volatility risk. In Section 4, I perform a set of robustness tests. In Section 5, I present the economic implications of my model. I conclude in the last section.

2 Model Development

Following Moise (2002), I test my hypothesis using a 2-factor model consisting of market return and innovations in its volatility. The two factors represent the myopic, and respectively intertemporal hedging components of asset demand. My model is tightly linked to Merton's ICAPM (1973) and provides a parsimonious and attractive way for testing the equilibrium asset pricing relationship. It basically says that expected returns should be equal to risk times the price for exposure to the risk inherent in the market portfolio and its volatility:

$$E_t[R^e_{i,t+1}] = \lambda_{m,t}\beta^m_{i,t} + \lambda_{IV,t}\beta^{IV}_{i,t} \qquad i = 1, 2, ..., n, \tag{1}$$

where

$$\beta^m_{i,t} = \frac{Cov_t\left(R^e_{i,t+1}, R^e_{m,t+1}\right)}{Var_t\left(R_{m,t+1}\right)} \quad and \quad \beta^{IV}_{i,t} = \frac{Cov_t\left(R^e_{i,t+1}, IV_{t+1}\right)}{Var_t\left(IV_{t+1}\right)}. \tag{2}$$

Here, R^e_i and R^e_m represent excess returns on asset i and the market portfolio, respectively. IV_t represents innovations in the realized stock market volatility and is equal to:

$$IV_t = v_{t,t+1} - v_{t-1,t}, \tag{3}$$

where $v_{t,t+\Delta}$ is computed by summing up the differences in squared daily market returns over the $[t, t+\Delta]$ time interval, with Δ representing the number of trading days in a given month, and then taking the square root of this quantity (see Anderson et al., 2003):

$$v_{t,t+\Delta} = \sqrt{\sum_{i=1}^{\Delta-1}(R_{m,t+i+1} - R_{m,t+i})^2}. \tag{4}$$

A plot of this monthly time series for the period January 1927-December 2005 can be seen in Figure 1[4].

In order to relate my study to existing literature (Ang et al., 2006 and Moise, 2002), I assume that the estimates in (2) are constant through time (thus, I take $\beta^m_{i,t} = \beta^m_i$ and $\beta^{IV}_{i,t} = \beta^{IV}_i$ $(\forall)$ $t = 1, ..., T$). Then, starting with the conditional model (1), I derive an equilibrium 2-factor asset pricing model for the unconditional expected returns:

[4] See French et al. (1987) for another measure of monthly stock market volatility.

$$E[R^e_{i,t+1}] = \Lambda_m \beta_i^m + \Lambda_{IV}\beta_i^{IV} \qquad i = 1, 2, ..., n, \tag{5}$$

where $\Lambda^m_{i,t}= E(\lambda_{m,t})$ and $\Lambda^{IV}_{i,t}= E(\lambda_{IV,t})$. Relation (5) says that an assets' risk can be determined by two estimates: the unconditional market beta and the unconditional innovations in volatility beta. This means that, when volatility is zero all assets have the same return as the risk-less asset, so the opportunity set reduces to that asset. When volatility is non-zero but non-volatile, the opportunity set is enlarged, although it does not change through time. Therefore, agents have more options, but no need for hedging. Furthermore, when volatility becomes volatile, it signals a change in the investment opportunities set and thus, it leads to agents' hedging demands.

First I estimate factors' loadings in Eq. (2). Previous empirical studies of financial markets found evidence in support for the disturbance variance in the time series models being less stable than usually assumed. Therefore, I estimate factors' loadings using an $AR(1) - GARCH(1,1)$ specification for the dynamics of the model error (Appendix A introduces the $GARCH$ models together with the stationarity and integrability conditions required for the models to work). This leads to standard errors which are corrected for autocorrelation and heteroskedasticity:

$$\left\{\begin{matrix} R^e_{i,t} = \alpha_i + \beta_i^m R^e_{m,t} + \beta_i^{IV} IV_t + \eta_{i,t} & \\ \eta_{i,t} = \varepsilon_{i,t} - \gamma_i \eta_{i,t-1}, & \\ \varepsilon_{i,t} = \sqrt{h_{i,t}} e_{i,t}, & i = 1, ..., 25, \\ h_{i,t} = \omega_i + \phi_{1,i} h_{i,t-1} + \theta_{1,i}\varepsilon^2_{i,t-1}, & \\ e_{i,t} \sim N(0,1). & \end{matrix}\right\} \tag{6}$$

In order to estimate system (6), I rely on the maximum likelihood method, which provides asymptotically efficient estimates[5].

Once the factors' loadings are estimated (the betas), I turn my attention towards estimating factors' risk prices in model (5) (the lambdas). In order to compare my results to previous findings, I employ the same procedure as previous authors did when estimating factors' risk prices, namely the Fama-MacBeth (1973) method[6]. Therefore, at each point in time I run a cross-sectional regression in order to estimate factors' risk prices. Then, I analyze the statistical distribution of the estimated risk prices. I report both

[5] The likelihood function is maximized via the dual quasi-Newton algorithm. The initial values for the regression parameters are obtained from the Yule-Walker estimates.

[6] Since returns are heteroskedastic, the usual approach is to employ the Generalized Method of Moments (GMM), which allows for efficient estimation in this case. When using this method, my results don't change.

the means and the t-statistics for testing whether the factors are priced in the cross-section of stock returns[7].

2.1 Economic Motivation

In order to study volatility effects on asset pricing, I present my model in the context of the standard theory of asset pricing in frictionless markets. Thus, the economic setting is a pure exchange economy (Lucas, 1978) with a single consumption good, no transaction costs, no short-sales constraints or any other market frictions, and a rational and infinitely lived representative consumer who has the recursive preferences proposed by Epstein and Zin (1989, 1991):

$$U_t = \left\{ (1-\delta)\, C_t^{\frac{1-\gamma}{\theta}} + \delta \left(E_t \left[U_{t+1}^{1-\gamma} \right]^{\frac{1}{\theta}} \right) \right\}^{\frac{\theta}{1-\gamma}}, \tag{7}$$

where

$$\theta = (1-\gamma)/(1-\psi^{-1}), \tag{8}$$

with C_t representing consumption at time t, $\psi > 0$ being the elasticity of intertemporal substitution, $\gamma > 0$ the coefficient of relative risk aversion, and $0 < \delta < 1$ the time discount factor. These preferences deviate from the power-utility model by relaxing the restriction that the elasticity of intertemporal substitution must equal to the reciprocal of the coefficient of relative risk aversion, while retaining the desirable scale-independence of the power utility function. The prices of securities are set in a general equilibrium model for the economy, where consumers choose the consumption and portfolio allocation that maximize their expected utility, while taking prices as given. The assumption of frictionless markets is combined with the assumption of no arbitrage, which implies the existence of a stochastic discount factor m_{t+1} that satisfies the dynamic Euler equation:

$$E_t \left[R^e_{i,t+1} m_{t+1} \right] = 0. \tag{9}$$

Earlier studies found little support for a consumption-based model (e.g., Hansen and Singleton, 1982, Mankiw and Shapiro, 1986). This may be due to the fact that aggregate consumption is a poor proxy for shareholder's

[7] In order to correct for the error-in-variables problem which is due to the fact that the betas are estimated and are not the "true" betas, the Shanken (1992) correction needs to be implemented here. When using it, the change in the results is very small.

consumption. Also, a return-based pricing kernel is a preferable alternative to a consumption-based pricing kernel, since consumption data is subject to substantial measurement error and periodic reevaluation. To avoid this problem, I follow Campbell (1993) who, using this economic setting, log-linearizes the Euler equation and substitutes consumption out of the asset pricing model, obtaining the following cross-sectional asset pricing formula:

$$E_t\left[r^e_{i,t+1}\right] = \gamma Cov_t\left(r^e_{i,t+1}, r^e_{m,t+1}\right) + (\gamma - 1) Cov_t\left(r^e_{i,t+1}, h_{t+1}\right) - \frac{1}{2} Var_t(r^e_{i,t+1}) \tag{10}$$

Equation (10) says that the expected excess return on an asset is determined by its own variance (a Jensen's inequality effect, which arises when log-returns are used instead of simple returns), and by a weighted average of two covariances. One is the asset's covariance with the market portfolio (the first term on the right-hand-side of relation (10)), while the other is the asset's covariance with news about future market returns (the second term in (10)). These two covariances represent the myopic and the intertemporal hedging components of asset demand, respectively. Thus, assets can be priced without direct reference to their covariance with consumption growth. Instead, covariances with the return on a portfolio of invested wealth, and with news about future returns on invested wealth are used. The latter covariance can also be thought of as the covariance with a hedging portfolio which indexes changes in the investment opportunity set. Please note that the elasticity of intertemporal substitution ψ is not present in Eq. (10) since consumption has been substituted out of the model.

So far, the asset pricing formula (10) is obtained under the assumptions of homoskedastic and jointly lognormal asset returns and consumption. Campbell (1993) relaxes these unrealistic assumptions and derives a new asset pricing formula which is similar to (10), with the exception that it also includes an extra term which reflects the sensitivity of consumption to the expected return on the market, through the precautionary savings channel. In this world, consumption can not be substituted out of the model with heteroskedastic returns. However, there are several special cases in which the substitution can be made. One of them is when the variance of the market returns follows a generalized autoregressive conditionally heteroskedastic ($GARCH$) process which is uncorrelated with the return on the market (Restoy, 1991). Since this correlation is very small in my data (-0.07), I follow Campbell and Vuolteenaho (2003) and derive an unconditional version of the first-order condition (10), by conditioning down and by replacing log returns with simple returns. I consider the log real risk-free rate as being approximately constant, an assumption which does not have a major influence on my tests, since I focus on stock portfolios. Therefore, I assume that the market return and the volatility processes are correlated with the stochastic discount factor, and I propose a pricing kernel m_{t+1} which is linear in these two variables:

$$m_{t+1} = \beta^m R^e_{m,t+1} + \beta^{IV} IV_{t+1.} \tag{11}$$

This leads to the specification of a parsimonious and testable model which makes no reference to consumption:

$$E\left[R^e_{i,t+1}\right] = \gamma Cov\left(R^e_{i,t+1}, R^e_{m,t+1}\right) + (\gamma - 1) Cov\left(R^e_{i,t+1}, h_{t+1}\right) \tag{12}$$

2.1.1 Market Volatility and Market Returns

In this section I argue that innovations in volatility plays the intertemporal hedging role, since a change in volatility signals a change in the investment opportunities set. To this end, I regress the news in market returns ($\Delta R_{m,t} = R_{m,t} - R_{m,t-1}$) on lagged IV:

$$\Delta R_{m,t} = \alpha + \beta IV_{t-1} + \varepsilon_t \tag{13}$$

My results indicate that the volatility factor is a good predictor for news about future market returns: $\hat{\beta} = 0.53$ ($t - stat = 6.45$) and $\hat{\alpha} = 0.00$ ($t - stat = 0.00$), with $R^2 = 4.2\%$.

Next I substitute (2) into Eq. (12) under the assumption that factors' loadings are time invariant. I obtain:

$$E\left[R^e_{i,t+1}\right] = \gamma \sigma^2_m \beta^m_i + (\gamma - 1) \sigma^2_{IV} \beta^{IV}_i, \tag{14}$$

where $\sigma^2_m = Var(R_{m,t+1})$ and $\sigma^2_{IV} = Var(IV_{t+1})$. Eqs. (5) and (14) lead to $\Lambda_m = \gamma \sigma^2_m$ and $\Lambda_{IV} = (\gamma - 1)\sigma^2_{IV}$.

Please remember that γ is the coefficient of relative risk aversion. Therefore, Eq. (14) captures the tight link that exists between volatility premium and investors' risk aversion[8] and shows the relevance of volatility risk for the cross-section of stock returns, which is also reflected by the intertemporal optimality condition (9), which holds for every asset i.

Please note that when the investment opportunity set is time-invariant (which translates into $Cov(R^e_{i,t+1}, h_{t+1}) = 0$) or when investors have unit

[8] Please note that industry practitioners have been using movements in implied volatility as an indicator for investors' risk aversion for some time.

coefficients of relative risk aversion ($\gamma = 1$) the multiperiod investment decision is just a repeat of identical one-period investment decisions. Therefore you only get the myopic demand, which is the first term on the right-hand side of equation (12).

3 Pricing Ex-Post Volatility Risk

The two testable implications of my model are the following. First, higher exposure to volatility leads to significantly lower average returns, which implies a negative volatility risk price. Second, there is a pattern of differential volatility loadings across the test assets matching the pattern in their average returns. Specifically, since it has been documented that small and value firms have historically higher average returns (the size and value anomalies), I should find higher volatility loadings for large and growth stocks.

In order to test these pricing implications I use monthly data, since relations in monthly returns have more economic significance and are less likely to be driven by frictions or data measurement issues, especially for large-firm returns. The time period under consideration is characterized by several recessions like the ones from 1933, 1945, 1949, 1954, 1958, 1961, 1981-1982, 1983, 1991, 2000-2001, major crises like the Penn Central commercial paper debacle of May 1970, the oil crisis of November 1973, the stock market crash of October 1987, the Asian crisis of 1997, and the Russian debt default of 1998 (see Figure 1). My test assets are the 25 size- and value-sorted portfolios of Fama and French (1992, 1993). I exclude the period July 1930-June 1931 due to unavailability of return data for some portfolios.

First, I analyze the time series of the risk factors' loadings. A caveat is in order. Volatility loadings are potentially time variable. Thus, the constant loadings constraint imposed in model (6) is not a natural one. For instance, if you write $\beta_{i,t}^{m} = \rho_{i,m}[\sigma_{i,t}/\sigma_{m,t}]$ and $\beta_{i,t}^{IV} = \rho_{i,IV}[\sigma_{i,t}/\sigma_{IV,t}]$, even if the idiosyncratic volatility is higher in times of high systematic volatility (or when volatility is more volatile), the correlation between portfolio returns and market returns (or the one between portfolio returns and innovations in aggregate volatility) is likely to be time variable. However, in order to compare my results with previous findings, I estimate constant factor loadings over the full sample.

For the 2-factor model to help in pricing asset returns, the intercepts from the time series regression should be statistically insignificant. Table I shows that this is the case mostly for the growth firms. Also, most of the volatility loadings and all of the market loadings are statistically significant.

My model explains at least 51% of the amount of variation in the time series of returns. From an economic perspective, the small growth portfolio is an outlier; its return is too low given its risk exposure. Also, value stocks have average returns that are too high given their risk exposure.

A careful examination of Table I and Figure 3 Panel b) reveals that, controlling for profitability, there is indeed a monotonically increasing pattern in volatility loadings as firm size increases. Also, controlling for the size of the firm, growth stocks tend to have higher volatility loadings than do value stocks. These results indicate a negative correlation between contemporaneous volatility loadings and average returns (see Figure 3 and Tables I and X). In the same time, Table I brings evidence in support of an $IGARCH(1,1)$ type of structure for the error variance (see Appendix B). Small portfolios exhibit a strong autoregressive pattern in their return series. The heteroskedasticity tests Portmanteau Q and Lagrange Multiplier LM strongly reject the null hypothesis of homoskedasticity in the error variance (results not reported). The time series plots of the conditional error standard deviations for the 25 portfolio returns also suggest that returns are heteroskedastic and that small stocks have the highest average conditional volatility (see Figure 4). As in Bollerslev (1987), the normality test indicates that the conditional normal distribution may not fully explain the leptokurtosis present in the portfolio returns (results not reported). In conclusion, even if there is not a significant temporal dependence in portfolio returns, heteroskedasticity and departure from normality in the time series of monthly portfolio returns pose a big problem.

Next, since volatility has a countercyclical pattern, risk-averse investors may require compensation for exposure to bad time risk, which is proxied here by innovations in volatility. Given the evidence that there is a positive spread in volatility loadings estimates between large and small firms, as well as between growth and value firms, which is negatively correlated with the corresponding spreads in average excess returns, I investigate whether volatility risk is priced in equilibrium. To this end, I perform unconditional tests of the asset pricing models[9]. Similarly to existing literature, I employ the Fama-MacBeth (1973) procedure and estimate time variable risk prices for my factors[10]. Table II reports the risk price estimates. Although there is on average a positive trade-off between risk and return over the time frame

[9] See Cochrane (2001) for a detailed discussion on deriving unconditional versions of asset pricing models.

[10] In a previous version of this paper I employed the GMM estimation technique, which is recommended in the presence of heteroskedasticity. The results are in the same direction with the findings documented here.

1927-2005 (significantly positive market risk price), a non-beta measure of risk like volatility plays a very important and apparently systematic role, having a noticeable influence on stocks' equilibrium expected returns. The sizeable variations in returns for large and growth firms are largely due to volatility risk exposure. I precisely estimate a negative volatility risk price. I infer that investors demand large and growth stocks since the small and value stocks expose them to larger downside risk. This pattern is consistent with a "flight-to-quality" argument: as recession deepens, credit markets get tighter and profitability decreases, and investors shift their preferences towards better collateralized, larger firms and towards more profitable, growth firms. In accordance with rational pricing, this hedging demand drives up the price for such stocks, which leads to lower average returns.

In Table III, I present monthly estimates for volatility premia earned by small stocks versus large stocks, for each value quintile. These premia depend on the dispersion in the size-related volatility loadings and on the average market price for volatility risk. Controlling for the accounting profitability type, the estimated yearly small-minus-big IV risk premia $\widehat{\lambda}_{IV}\left(\widehat{\beta}^{IV}_{Small} - \widehat{\beta}^{IV}_{Big}\right)$ ranges from 3% to 7% . You can also see that the volatility premium is larger for value versus growth firms. I conclude that differences in estimated volatility loadings are sufficient to account for differences in returns for common stocks of firms of different sizes and different accounting profitability type.

Please note that the simple net returns for the 25 portfolios are log-Normally distributed. You need to be careful when interpreting the economic significance of the lambda estimates in the above model. What you may want to do is to adjust the risk premia using the expected value of the log-Normal distribution, $E(R^e_{i,t}) = \exp(\sigma^2_{i,t}/2)$. Here, the $\sigma^2_{i,t}s$ represent individual portfolios' conditional variances. Using the notation from model (6), the implied unconditional variance (the long-run variance) for each portfolio is expressed as $\sigma^2_i = \hat{\omega}_i/(1 - \hat{\phi}_i - \hat{\theta}_i)$. Therefore, one can derive individual long-run risk premia earned by the size- and value- sorted portfolios by using these relations.

4 Robustness Checks

In this section I ensure that my results are robust with respect to the volatility measure employed, to the inclusion of classical risk factors, to different model specifications, and that they hold for different samples.

4.1 Alternative Volatility Measures

4.1.1 GARCH Models

There are a few stylized facts about stock return volatility that have been brought to light: it increases after a drop in stock prices (Black, 1976, French, Schwert, and Stambaugh, 1987, Nelson, 1990), it is persistent (Poterba and Summers, 1986, Schwert, 1989, French, Schwert, and Stambaugh, 1987), and it is related to macroeconomic volatility and recessions (Officer, 1973, Pindyck, 1984, Schwert, 1989, 1990, Hamilton and Lin, 1996). An underlying model for volatility which captures these documented features needs to be built in order to predict volatility. For this purpose it is important to remember that previous studies of financial markets found evidence in support of the disturbance variance in the time series models being less stable than usually assumed. This result is due to the fact that large and small forecast errors appear to occur in clusters, suggesting a heteroskedastic pattern in the variance of the forecast error, which depends on the size of the previous disturbance. Small price changes tend to be followed by small price changes, and vice versa (Mandelbrot, 1963). Thus, the first measure of volatility used here although simple, ignores time-series properties like volatility clustering (which is just a clustering of information arrivals), autoregression and heteroskedasticity, all of which are present in real data. More sophisticated statistical models are needed in order to capture the time-variation in returns. Simple filters, such as the rolling standard deviation used by Officer (1973), have given way to parametric *ARCH* (*Autoregressive Conditionally Heteroskedastic*) or stochastic-volatility models. Engle (1982) introduces *ARCH* models for describing stochastic volatility, which prove to be helpful in analyzing financial time series. The *ARCH* model has, among other problems, the weaknesses of assuming symmetric effects of return shocks on volatility and a big set of parameters needed for describing volatility. A better alternative for modeling persistent movements in volatility, more parsimonious and also more flexible, has been introduced by Bollerslev in 1986 and it is called *GARCH* (*or Generalized ARCH*)(see Appendix A for a description of the GARCH models)[11].

In this study I model the volatility of market returns using a parsimonious model, which is representative for the GARCH family and it is widely

[11] Partial surveys of the big literature on these models can be found in Bollerslev et al. (1992), Hentschel (1995), Ghysels et al. (1996), Campbell et al. (1997, Chap. 12) and Tsay (2001).

used in finance: a $GARCH(1,1)$ model. This model is conditional on the last period, with previous volatility encompassing the set of all relevant information, and it is feasible and competitive in analyzing financial time series. The heteroskedastic characteristic of $GARCH$ models explains some of the excess kurtosis present in returns time series, but usually not all of it. Fat tail distributions like the $T-distribution$ are needed to enhance the results. Engle and Bollerslev (1986) introduce a $Student-distribution$ model meant to capture this phenomenon. The fat tails of the *Student distribution* allow it to capture the kurtosis present in the financial data better than the $Normal\ density$ (Bollerslev, 1992). Thus, I employ an $Asymmetric-Student\ GARCH\ (1,1)$ model for the market portfolio returns series.

There are two implications of the $GARCH$ models that need to be addressed first. The first one is that the observed series has a constant mean of zero. However, Campbell et al. (1997) point out that a positive autocorrelation of order one is usually found in an index' return series. For this reason, an $AR(1)$ model is specified for the stock market return series. The second implication is that the model is symmetric in that both positive and negative shocks have the same impact on volatility (squared returns drive revisions in the forecasts). In other words, the plot of conditional volatility on the past shock, also called the news impact curve, is symmetrical about 0. However Black (1976), among others, pointed out an asymmetry in the stock market data: negative innovations in stock returns have a greater impact on volatility than do positive ones. One explanation for this phenomenon is the leverage hypothesis, posited originally by Black (1976), which says that negative shocks to returns drive up volatility: a decline in stock prices increases leverage (since the value of the equity becomes a smaller share of the total), increases the expected return on the stock, and increases the variance of stock's return. The other explanation is based on the volatility-feedback hypothesis (Campbell and Hentschel, 1992), which says that positive shocks to volatility drive down returns: if expected stock returns increase when volatility increases and if expected dividends are unchanged, then stock prices should fall. In response to these issues, the *exponential* $GARCH$ model of Nelson (1991), the $ARCH$ model of Zakoian (1994), and the GJR *model* of Glosten, Jagannathan and Runkle (1993) brought in several specifications designed to introduce asymmetry into the model. In this study I employ the $GJR-type\ Asymmetric\ Student-GARCH$ model. Figure 2 shows the time series plot of the index return data that I use to predict volatility.

4.1.2 Estimation

$GARCH$ models characterize the conditional distribution of returns innovations by imposing serial dependence on their conditional variance. Large disturbances, positive or negative, become part of the information set used to construct the variance forecast for the next period's disturbance. Thus, large shocks are allowed to persist, capturing the volatility clustering phenomenon. As described above, I estimate a $GJR-type\ Asymmetric\ Student-GARCH$ model for volatility using monthly returns on the market portfolio:

$$\left\{\begin{array}{cc} R_{m,t} = \mu + \rho R_{m,t-1} + \eta_t & \\ \eta_t = \varepsilon_t \sqrt{h_t}, & \\ \varepsilon_t / I_{t-1} \sim Student\ (\upsilon), & t = 1, ..., T. \\ h_t = \alpha + \phi h_{t-1} + \theta^{+} \eta_{t-1}^{2+} + \theta^{-} \eta_{t-1}^{2-}, & \\ \eta_t^{2+} = \eta_t^2 1_{\{\eta_t > 0\}},\ \eta_t^{2-} = \eta_t^2 1_{\{\eta_t < 0\}} & \end{array}\right\} \tag{15}$$

This model accommodates both the asymmetry in the news impact curve and the fat-tail behavior of stock returns and allows for non-zero first order correlation. The degrees of freedom variable for the Student-t distribution is meant to capture the excess kurtosis present in the index return data. Following Bauwens and Lubrano (1998), I consider a Half-Cauchy prior for υ and flat priors on finite intervals for all the other parameters. I use a $Griddy-Gibbs\ sampler$ as described by Ritter and Tanner (1992) applied to bivariate posterior densities in order to estimate this model (see Appendix C for a description of the Griddy-Gibbs sampler). The *Gibbs sampler* requires analytical knowledge of the full conditional posterior densities. Regression models with $GARCH$ errors do not contain this knowledge. To handle this, I apply a unidimensional deterministic integration rule to each coordinate of the posterior density in combination with the *Gibbs sampler*, as described by Bauwens and Lubrano (1998). The random draws of the joint posterior are then obtained by evaluating and inverting the full conditional densities. Please note that the $Griddy-Gibbs\ method$ proves useful in cases involving high dimensional posterior densities.

Now I estimate model (15). The results are computed using a $Griddy-Gibbs\ sampling$ algorithm, in which 1000 draws are kept and 1000 initial draws are the burn-ins sample. The posterior results are reported in Table VIII. Similar to the findings of Schwert (2002) for the NASDAQ and S&P composite portfolio, I find that conditional volatility is persistent over time: the $GARCH$ parameter ϕ is precisely estimated to be equal to 0.75. The extent to which a volatility shock today feeds through into next period's volatility is equal to 0.75. Notice that the stationarity constraint (A3) is satisfied close to the boundary, since the sum $\phi+\theta^{-} = 0.75+0.18 = 0.95$. I find evidence of volatility feedback effect in my sample, since only the negative shocks coefficient estimate is precisely estimated. The leverage hypothesis of Black (1976) is also supported by my results: the volatility feedback effect together with the persistence in volatility lead to a negative correlation between current market returns and future volatility.

4.2 Pricing Ex-Ante Volatility

In an efficient capital market, when making their asset allocation, investors use the best conditional forecasts of variables like market volatility, since they affect equilibrium expected returns. While market returns are hard to predict, market volatility is predictable (Bollerslev et al., 1992). In order to build ex-ante volatility I use the posterior estimates of model (15) (Table IX). Figure 1 plots this series next to the ex-post volatility series. You can notice that they track each other closely, result due to the fact that the $GARCH$ predicting equation is in fact a weighted average of past squared returns, with slowly declining weights.

By repeating the analysis using the new volatility factor, I find that the price of volatility risk for ex-ante volatility is also precisely estimated and negative (see Table VII). Therefore, my findings are robust with respect to the type of volatility measure employed. Please note that the previous measure, ex-post volatility, includes both the ex-ante volatility and the unexpected changes in volatility, which justifies the difference in the estimated risk premia between ex-post and ex-ante volatility.

4.3 Other Risk Factors

Fama and French (1996) rely on their 3-factor model for explaining return anomalies related to firm characteristics. In order to quantify the effect of volatility risk over and above the effects of these factors, the 2-factor asset pricing model (1) is augmented with the HML and SMB factors of Fama and French[12]. While the SMB risk price is sometimes imprecisely estimated in my sample, the market, volatility and HML risk prices are precisely estimated (Table IV). The volatility effect has a similar magnitude to the one obtained when using the 2-factor model.

[12] HML and SMB are mimicking portfolios for book-to-market equity and size (zero-investment portfolios). HML is the difference between high book-to-market-stocks portfolios and low book-to-market-stocks portfolios, with similar weighted-average size, while SMB is the difference between the returns on small-stocks portfolios and those of big-stocks portfolios, with similar weighted-average book-to-market equity.

Another variable you would want to control for is the momentum factor of Jegadeesh and Titman (1993). When the previous 4-factor model is augmented with this risk factor, the price implications for volatility risk remain unchanged, while the momentum factor is imprecisely estimated. These results are consistent with volatility risk being priced and they also show that the volatility effect is not subsumed by traditional risk factors (factors' cross-correlation matrix is presented in Table X Panel b).

I document the empirical failure of $CAPM$ in Figure 5 Panel a). A more stringent test consists of assessing the empirical performance of my 2-factor model versus the FF-3 factor model. The 2-factor model with an intercept performs almost as well as the Fama-French 3-factor model in explaining the cross-sectional variation in stock returns (see Figure 5 Panels b and d). The slightly superior performance of the latter model in terms of *average* MSE is not surprising, given that the SMB and HML factors are so successful in capturing documented empirical patterns in returns.

4.4 Different Model Specifications

Now I assume that the borrowing and lending rates in the market are different, which leads to estimating a constant in the cross-sectional model (1). The implications from allowing the constant in the model are negligible for the volatility, SMB, HML and MOM factors, but they are dramatic for the market factor, since they lead to results that are hard to believe: a significant and negative market risk price, which implies a negative market risk premium (see Table VI). The reason for this is that there is a measurement error in model (1): the estimated covariances between returns and risk factors are generated regressors and not "true" betas (the "errors-in-variables" problem). They are noisy and may lead to a significant intercept in the cross-sectional regression. If the factors form a basis for the space of test assets, and if the factors are traded in the market, then their risk prices should be close in value to their means. Table IV and Table X Panel a) show that this is the case when using the model which assumes that the lending and borrowing rates are identical (or equivalently, $\alpha = 0$). Nonetheless, the cross-sectional regression that assumes that the two rates are different (or equivalently, $\alpha \neq 0$) leads to a better fit of predicted versus realized average excess returns (see Figure 5, Panels b and c). In conclusion, estimating a constant in the cross-sectional model can make a big difference, since freeing up the constant allows the regression line to better fit the data.

Also, in order to compare the relative performance of the different asset pricing models employed here, similarly to Jagannathan and Wang

(1996), I use the R^2 from the cross-sectional regression of average excess returns on the risk factors to measure the amount of variation in cross-sectional returns explained by the models. Since the coefficient of multiple determination, R^2, can be made large by including a large number of independent variables, I also report the adjusted coefficient, R^2_{adj}, which adjusts for the number of factors in the models.

4.5 Different Samples

Now I test my findings using two additional samples: the post-depression sample (January 1935-December 2005) and the post-COMPUSTAT sample (January 1963-December 2005). Results are in the same direction (see Table V). Please note that the MOM factor becomes significant in the post-COMPUSTAT sample. Also, volatility has a smaller price of risk in this sample, result probably due to the fact that there are fewer recessions during this period, which are very important for detecting downside risk.

5 Economic Implications

The evidence I present here implies that in the $ICAPM$ world the market portfolio is not mean-variance efficient with respect to the universe of common stocks, and suggests adding a position which takes into account fluctuations in volatility, which is a hedging instrument. Since investors are not fully insured against systematic volatility risk, the size- and value-related premia reflect their attempts to reduce this risk exposure. By using volatility as a hedge, they insure themselves against decreases in the Sharpe Ratio of their wealth portfolio.

My results also have a clear economic interpretation based on the tight link between a representative consumer's coefficient of relative risk aversion and the volatility risk premium, as outlined in Section 2.1.1. The hedging demand of a representative consumer varies with γ. Since $\hat{\lambda}_{IV}$ is equal to $(\gamma - 1)\sigma^2_{IV}$, I can infer from Table II that the implied coefficient of relative risk aversion is less than 1, which is well below the maximum value of 10 considered plausible by Mehra and Prescott (1985). Thus, the representative consumer with recursive preferences outlined in model (7) is risk averse, although less than the log-utility maximizing agent. He is willing to accept on average lower returns in order to hold stocks that pay off when wealth is most productive, which is a key ingredient of the risk premia observed in the stock market.

6 Conclusions

In this study I show that stock market volatility changes through time in a stochastic and fairly persistent fashion which leads to sizeable capital gains or losses in the universe of common stocks. Exposure to volatility risk is of economic importance and helps explain a sizable variation in the cross-sectional return heterogeneity. More interestingly, my results have broader implications. Besides establishing that volatility is a systematic source of risk negatively priced in the long run in the market, I also document a link between volatility risk and the premia associated with firm characteristics. This delivers additional insights into some of the economic fundamentals lying behind the size and book-to-market anomalies, and supports the concept of market efficiency. I infer that large stocks are perceived by investors as a means of intertemporal hedge. Their returns co-vary positively with volatility. Therefore, they pay off during bad economic times, when volatility is more volatile and market returns are lower. In accordance with rational pricing, this leads to higher hedging demands, higher prices and lower average returns for large stocks. I find evidence that growth stocks have hedging value for investors, too. I conclude that investors are willing to forgo expected returns in order to get downside protection. This is in agreement with a "flight-to-quality" during bad economic times. Investors shift their preferences towards better collateralized, larger firms and towards more profitable, growth firms. Controlling for accounting profitability type, I estimate an yearly volatility risk premium ranging from 3% to 7%. On average, I find that over 90% of the size anomaly is explained by exposure to volatility risk. I also find that the volatility effect is not subsumed by any of the size, value or momentum effects, that it is robust with respect to the type of volatility measure used, that it is not sample specific or model dependent.

There are some practical difficulties when building a hedging strategy based on the results of this study. The hedging strategy is subjected to the difficulty in assessing next period volatility. When making the portfolio allocation, volatility can not be observed directly, and thus has to be estimated using historical data, which adds measurement error to the forecast. If the volatility estimate is incorrect, so will be the hedging position. Nonetheless, the $Asymmetric - GARCH$ model presented here is a good starting point, since it captures features like persistence and asymmetry in conditional volatility, pointing to the volatility feedback effect and the leverage hypothesis of Black (1976). Also, since the market portfolio proxy may not be a very good one, its volatility can be improved upon, too.

I believe that my efforts have revealed sufficient economic significance to justify continued research on asset pricing models incorporating aggregate volatility risk. Volatility plays a leading role in helping investors and policy makers better understand the market. My 2-factor model has a strong link to $ICAPM$ and looks manageable both in an academic and industrial environment. Since imposing constant factor loadings may not be a natural assumption, I plan to turn my attention towards an asset pricing model which links equity premia to time variable factor loadings. In addition, I plan to investigate the effect aggregate volatility has in an international framework, by hinging on the relation between volatility and country-specific business cycles.

7 Appendix A

GARCH (1, 1) is written as:

$$\left\{ \begin{array}{c} R_t = \mu + \varepsilon_t \\ \varepsilon_t / I_{t-1} \sim N(0, h_t) \\ h_t = \alpha + \phi h_{t-1} + \theta \varepsilon_{t-1}^2 \end{array} \quad , \ t = 1, ..., T \right\} \tag{16}$$

The distribution of R_t is Normal with mean μ and variance h_t given past information I_{t-1}. I assume the ε_t sequence to be uncorrelated, but not independent. Nonetheless, the standardized disturbance, $\varepsilon_t / h_t^{1/2}$, is *iid* $N(0, 1)$. The initial variance is a known constant. The parameters ϕ and θ are restricted to ensure that the conditional variance is positive:

$$\alpha > 0, \phi \geq 0, \theta \geq 0. \tag{17}$$

To render the covariance stationarity of y_t, the following restriction is imposed:

$$0 \leq \phi + \theta < 1. \tag{18}$$

The y_t process can be strongly stationary even if it is not covariance stationary because the condition for strong stationarity,

$$E[\log(\phi \varepsilon_t^2 + \theta) < 0 \tag{19}$$

is weaker than the condition for weak stationarity (Nelson, 1990), which ensures the existence of unconditional variance. So, a necessary condition for any stationarity is $\phi < 1$.

$GARCH(1, 1)$ with student errors, or $Student - GARCH$ is written as:

$$\left\{ \begin{array}{c} R_t = \varepsilon_t \sqrt{h_t} \\ \varepsilon_t / I_{t-1} \sim Student(\upsilon) \\ h_t = \alpha + \phi h_{t-1} + \theta \varepsilon_{t-1}^2 \end{array} \quad , \ t = 1, ..., T \right\} \tag{20}$$

The distribution of R_t is *Student* with mean zero and variance $h_t^{1/2} \upsilon / (\upsilon - 2)$ given past information I_{t-1} and assuming $\upsilon > 2$. Again, the ε_t sequence is independent and the initial variance is a known constant. Condition (18) needs to hold again to ensure that the conditional variance is positive. In this case (19), the covariance stationarity condition, becomes:

$$0 \leq \theta \frac{\upsilon}{\upsilon - 2} \phi < 1 \quad (\upsilon > 2). \tag{21}$$

Again, a necessary condition for any stationarity is ϕ <1.

Now let γ denote the parameter vector $(\alpha, \theta, \phi, \upsilon)$. The posterior density for a sample of T observations is given by

$$\varphi(\gamma/R) \propto \varphi(\gamma) l(\gamma/R), \tag{22}$$

with the likelihood function given by

$$l(\gamma/R) \propto \prod_{t=1}^{T} \frac{\Gamma(\frac{\upsilon+1}{2})}{\Gamma(\frac{\upsilon}{2})} (\upsilon h_t^{1/2})^{-1/2} \left[1 + \frac{R_t^2}{\upsilon h_t^{1/2}}\right]^{-\frac{\upsilon+1}{\upsilon}} \tag{23}$$

where the prior density, $\varphi(\gamma)$, needs to respect the positivity restrictions on the parameters and the condition $\phi < 1$. Integrability of the posterior density depends in part on the integrability of the prior density. Given an integrable (or proper) prior and a non-pathological likelihood, the posterior will also be integrable. Examining the likelihood function (23) it is found that, if $h_t^{1/2}$ is strictly positive, since the *Student* density is finite and positive, no pathology appears. Therefore, a flat prior (like the uniform) may be used for these parameters. However, the posterior density of υ is not integrable when using a flat prior (see Bauwens and Lubrano, 1998). For the posterior density of υ to be integrable, the prior information must be such that the posterior is forced to go to zero quickly enough in the tail. The prior at the right tail should be at least $O(\upsilon^{1+d})$, with d being small and positive, e.g. $1/\upsilon^2$ (improper prior obtained by being flat on $1/\upsilon$). This prior must be truncated to the interval (m, ∞), with m being small and positive, to avoid causing problems at the left tail. This avoids the problem of $l(\gamma/R)/\upsilon^2$ approaching infinity as υ approaches zero. For a proper prior for υ, a $half - right\ Cauchy$ centered at 0 is used:

$$\varphi(\upsilon) \propto (1 + \upsilon^2)^{-1} \quad (\upsilon > 0). \tag{24}$$

Some of the other possibilities for the prior on υ may include a flat prior on υ over a finite range $(0, M)$, and also an exponential density (Geweke, 1993), which uses a subjective parameter chosen to fix the prior mean and variance of υ (see Bauwens and Lubrano, 1998).

8 Appendix B

The *GARCH* model is estimated using the maximum likelihood method. The log-likelihood function is computed from the product of all conditional densities of the prediction error:

$$l = \sum_{t=1}^{T} \frac{1}{2} \left[-\ln(2\pi) - \ln(h_t) - \frac{\varepsilon_t^2}{h_t} \right] \tag{25}$$

The condition

$$\sum_{i=1}^{q} \theta_i + \sum_{j=1}^{p} \varphi_j < 1 \tag{26}$$

implies that the *GARCH* process is weakly stationery. In the presence of autocorrelation, the stationarity condition is

$$\frac{1}{1-\gamma^2} \sum_{i=1}^{q} \theta_i + \sum_{j=1}^{p} \varphi_j < 1 \tag{27}$$

When the model is integrated in variance (*IGARCH*),

$$\sum_{i=1}^{q} \theta_i + \sum_{j=1}^{p} \varphi_j = 1 \tag{28}$$

The interesting feature of *IGARCH* models is that they are strongly stationary, although not weakly stationary.

9 Appendix C

The *Gibbs sampler* of Geman and Geman (1984) and Gelfand and Smith (1990) is a very popular $MCMC$ method. Let $\theta_1,\ \theta_2,\ ...\theta_n$ be a set of parameters that need to be estimated, X the available data, and M the model entertained. Suppose that the conditional distributions of each parameter given the others, $f_i(\theta_i/\theta_{j\neq i}, X, M)$ are known, but the likelihood function of the model is hard to obtain. What I do is to draw a random number from each of these conditional distributions. For instance, if $n = 3$, let's consider $\theta_{2,0}$and $\theta_{3,0}$ two arbitrary starting values of θ_2 and θ_3. Then

1. A random sample $f_1(\theta_1/\theta_{2,0}, \theta_{3,0}, X, M)$ is drawn, call it $\theta_{1,1}$;
2. A random sample $f_2(\theta_2/\theta_{3,0}, \theta_{1,1}, X, M)$ is drawn, call it $\theta_{2,1}$;
3. A random sample $f_3(\theta_3/\theta_{2,1}, \theta_{1,1}, X, M)$ is drawn, call it $\theta_{3,1}$.

This is a *Gibbs iteration*. The iteration can be repeated for n times, with n sufficiently large such that $m < n$ initial random draws can be discarded. I get the *Gibbs sample* this way, $(\theta_{1,m+1}, \theta_{2,m+1}, \theta_{3,m+1}), ...(\theta_{1,n}, \theta_{2,n}, \theta_{3,n})$,which can be used to obtain the point estimates and the variances of the three parameters.

In the case when the conditional posterior distributions of the parameters don't have closed-form expressions, the *Gibbs sampler* implementation can become complicated. But Ritter and Tanner (1992) have a method to obtain draws in this case. It is called the $Griddy - Gibbs\ sampler$:

1. A grid of points are chosen from a properly selected interval of θ_i, say $\theta_{i1} \leq \theta_{i2} \leq ... \leq \theta_{im}$. The conditional posterior density function is evaluated to obtain $w_j = f(\theta_{ij}/\theta_{lk\neq ij}, X, M)$ for $j = 1, ...m$;
2. $w_1, ..., w_m$ are used to obtain an approximation to the inverse cumulative distribution function of $f(\theta_{ij}/\theta_{lk\neq ij}, X, M)$;
3. A $Uniform(0, 1)$ random variate is drawn and the observation is transformed via the approximate inverse CDF in order to obtain a random draw for θ_i.

The usual *Gibbs sampler* cannot be applied to the $GARCH$ model even if the error term is (conditionally) normal. Instead, the $Griddy - Gibbs$ *sampler* can be applied to bivariate posterior densities. The $Griddy-Gibbs$ *method* proves useful in cases involving high dimensional posterior densities, such as the seven-parameter model shown in this study.volatility and country-specific business cycles.

References

[1] Anderson, T.G., T. Bollerslev, F.X. Diebold and P. Labys, 2003. Modeling and Forecasting Realized Volatility. Econometrica.

[2] Ang, A., R.J. Hodrick, Y. Xing, Y. and X. Zhang, 2006. The Cross-Section of Volatility and Expected Returns. Journal of Finance.

[3] Bauwens, L. and M. Lubrano,1998. Bayesian Inference on GARCH Models Using the GibbsSampler. Econometrics Journal, 1, C23-C46.

[4] Bernanke, B. and M. Gertler, 1989. Agency costs, net worth, and business fluctuations. American Economic Review 79, 14-31.

[5] Black, F., 1972. Capital Market equilibrium with Restricted Borrowing. Journal of Business 45, 444-455.

[6] Black, F., 1976. Studies of Stock Price Volatility Changes. Proc. Am. Stat. Assoc., Bus. Econ. Stat. Section, 177-81.

[7] Black, F., C. Jensen and M. Scholes, 1972. The Capital Asset Pricing Model: Some Empirical Tests. Studies in the Theory of Capital Markets.

[8] Black, F. and M. Scholes, 1973. The Pricing of Options and Corporate Liabilities. Journal of Political Economy 81, 637-654.

[9] Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307-27.

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

[11] Bollerslev, T., R.Y. Chou and K.F. Kroner, 1992. ARCH Modeling in Finance. Journal of Econometrics 52,5-59.

[12] Bollerslev, T. and R.F. Engle, 1986. Modeling the Persistence of Conditional Variances. Econometric Rev.5, 1-50.

[13] Bollerslev, T., M. Gibson, M. and H. Zhou, 2004. Dynamic Estimation of Volatility Risk Premiaand Investor Risk Aversion from Option-Implied and Realized Volatilities. Working Paper, Federal Reserve Board, Washington, D.C.

[14] Campbell, J.Y., 1993. Intertemporal Asset Pricing without Consumption Data. The American Economic Review, Vol.83, No.3, 487-512.

[15] Campbell, J.Y. and L. Hentschel, 1992. No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns. Journal of Financial Economics 13, 281-318.

[16] Campbell, J.Y., W. Lo and A.C. Mac Kinley, 1997. The Econometrics of Financial Markets. Princeton: Princeton University Press.

[17] Campbell, J.Y. and T. Vuolteenaho, 2004. Bad Beta, Good Beta. Forthcoming, American Economic Review.

[18] Chen, J., 2002. Intertemporal CAPM and the Cross-Section of Stock Returns. Working Paper, University of Southern California.

[19] Cochrane, J., 2001. Asset Pricing. Princeton University Press.

[20] Engle, R.F., 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50, 987-1007.

[21] Epstein, L. and S. Zin, S. (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. Econometrica 57, 937-969.

[22] Epstein, L. and S. Zin, S. (1991). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Investigation. Journal of Political Economy, 99, 263-286.

[23] Fama, E. F. and K.R. French, 1992. The Cross-Section of Expected Stock Returns. Journal of Finance 47, 427-465.

[24] Fama, E. F. and K.R. French, 1993. Common Risk Factors in the Returns on Stocks and Bonds. Journal of Financial Economics 33, 3-56.

[25] Fama, E. F. and K.R. French, 1995. Size and Book-to-Market Factors in Earnings and Returns. Journal of Finance 50, No. 1, 131-155.

[26] Fama, E. F. and K.R. French, 1996. Multifactor Explanations of Asset Pricing Anomalies. Journal of Finance 60, No. 1, 396-425.

[27] Fama, E. F. and J. MacBeth, 1973. Risk, Return, and Equilibrium: Empirical Tests. Journal of Political Economy, 71, 607-636.

[28] French, K.R., G.W. Schwert, G.W. and R.F. Stambaugh, 1987. Expected Stock Returns and Volatility. Journal of Financial Economics, 19, 3-30.

[29] Gallant, A.R., 1987. Nonlinear Statistical Models. New York: John Wiley and Sons, Inc.

[30] Geweke, J., 1993. Bayesian Treatment of the Independent Student-T Linear Model. Journal of Applied Econometrics 8 (Supplement), 19-40.

[31] Ghysels, E., A. Harvey, and E. Renault, 1996. Stochastic Volatility. Handbook of Statistics, Vol. 14, G.S.Maddala, ed., North-Holland, Amsterdam.

[32] Glosten, L. R., R. Jagannathan and D.E. Runkle, 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. Journal of Finance 48, 1779-801.

[33] Hamilton, J., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

[34] Hamilton. J. D. and G. Lin, 1996. Stock Market Volatility and the Business Cycle. Journal of Applied Econometrics, Vol. 11, No. 5, 573-593.

[35] Hentschel, L., 1995. All in the Family: Nesting Symmetric and Asymmetric GARCH Models. Journal of Financial Economics, 39, 71-104.

[36] Hull, J.C., 1999. Options, Futures, and Other Derivatives. Fourth Edition, Englewood Cliffs,NJ:Prentice Hall.

[37] Jacquier, E., N.G. Polson and P.E. Rossi, 1994. Bayesian Analysis of Stochastic Volatility Models. Journal of Business and Economic Statistics 12, 371-389.

[38] Jacquier, E., N.G. Polson and P.E. Rossi, 2004. Bayesian Analysis of Stochastic Volatility Models With Fat-Tails and Correlated Errors. Journal of Econometrics, 122(1), 185-212.

[39] Jagannathan, R. and Z. Wang, 1996. The Conditional CAPM and the Cross-Section of Expected Returns. Journal of Finance 51, 3-53.

[40] Jegadeesh, N. and S. Titman, 1993. Returns to Buying Winners and Selling Losers: Implicationsfor the Stock Market Efficiency. Journal of Finance 48, No. 1, 65-92.

[41] Lucas, R.E., Jr., 1978. Asset Prices in an Exchange Economy. Econometrica, 46, 1429-1445.

[42] Mandelbrot, B., 1963. The Variation of Certain Speculative Prices. Journal of Business, 36, 394-419.

[43] Mehra, R. and E. Prescott, 1985. The Equity Premium: A Puzzle. Journal of Monetary Economics,15, 145-161.

[44] Merton, R.C., 1973. An Intertemporal Asset Pricing Model. Econometrica, 41, 867-887.

[45] Merton, R.C., 1990. Continuous-Time Finance. Basil Blackwell Inc., UK: T J Press Ltd.

[46] Moise, C.E., 2002. Is Market Volatility Priced? Working Paper, University of Chicago.

[47] Nelson, D.B., 1990. Stationarity and Persistence in GARCH (1, 1) Model. Econometric Theory 6, 318-334.

[48] Nelson, D.B., 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. Econometrica, 59, 347-370.

[49] Officer, R.R., 1973. The Variability of the Market Factor of New York Stock Exchange. Journal of Business 46, 434-453.

[50] Perez-Quiros, G. and A. Timmermann, 2000. Firm Size and Cyclical Variations in Stock Returns.Journal of Finance 55, No.3, 1229-1262.

[51] Petkova, R. and Zhang, L, 2005. Is Value Riskier than Growth? Journal of Financial Economics 78, 187-202.

[52] Restoy, F., 1991. Optimal Portfolio Policies under Time-Dependent Returns. Unpublished Manuscript, Harvard University.

[53] Ritter, C. and M.A. Tanner, 1992. Facilitating the Gibbs Sampler: the Gibbs Stopper and theGriddy-Gibbs Sampler. Journal of Am. Stat. Assoc. 87, 861-8.

[54] Schwert, G.W., 1989. Why Does Stock Market Volatility Change Over Time? Journal of Finance 44, No 5, 1115-1153.

[55] Schwert, G.W., 1990. Stock Volatility. Financial Analysts Journal, May-June 1990.

[56] Schwert, G.W., 2002. Stock Volatility in the New Millennium: How Wacky Is NASDAQ? Journal of Monetary Economics 49, 3-26.

[57] Shanken, J., 1992. On the Estimation of Beta Pricing Models. Review of Financial Studies 5, 1-34.

[58] Tsay, R., 2001. Analysis of Financial Time Series. John Wiley and Sons, Inc.

[59] Zakoian, J-M., 1994. Threshold Heteroskedastic Models. Journal Econ. Dyn. Control 18, 931-55.

[60] Vayanos, D., 2004. Flight to Quality, Flight to Liquidity, and the Pricing of Risk. Working Paper, NBER.

Figure 1
Ex-Post and Ex-Ante Stock Market Volatilities

Ex-post volatility is built on a monthly basis for the period January 1927 – December 2005, using NYSE, AMEX, and NASDAQ return files maintained by CRSP. It is similar to the one used by Anderson et al. (2002), and is given by

$$v_{t,t+\Delta} = \sqrt{\sum_{i=1}^{\Delta-1}\left(R_{m,t+i+1} - R_{m,t+i}\right)^2}$$

where $R_{m,t}$ denotes daily return on the stock market portfolio and Δ represents the number of trading days in a given month. The ex-ante volatility series is computed using an *Asymmetric-Student GARCH (1,1)* model:

$$\begin{cases} \eta_t = R_t - 0.201 - 0.047 R_{t-1}, \\ h_t = 0.088 + 0.025\eta_{t-1}^{2+} + 0.182\eta_{t-1}^{2-} + 0.753 h_{t-1} \end{cases}.$$

Both series are reported at a monthly level. The ex-post volatility series is slightly larger than the ex-ante series. For each year, the tick marks correspond to the month of January. The shaded areas represent NBER recessions or financial distress time periods.

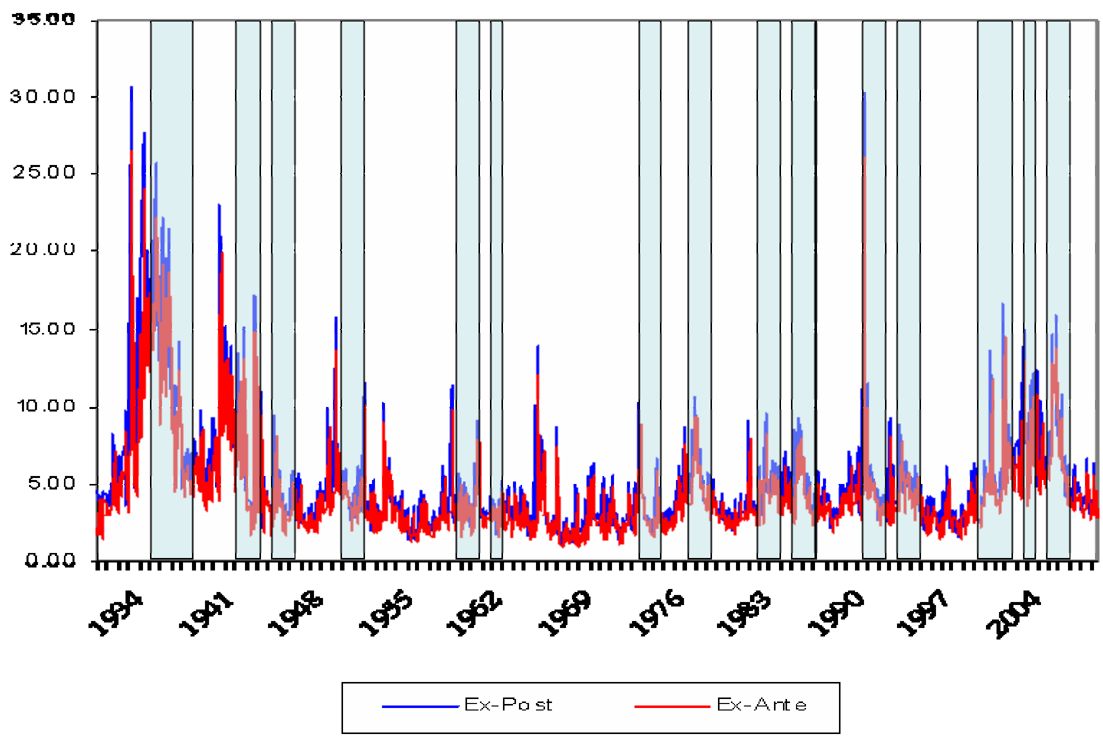


Figure 2
Stock Market Index Excess Returns

The monthly time series of excess returns for the stock market index is presented for the period January 1927 – December 2005. For each year, the tick marks correspond to the month of January.

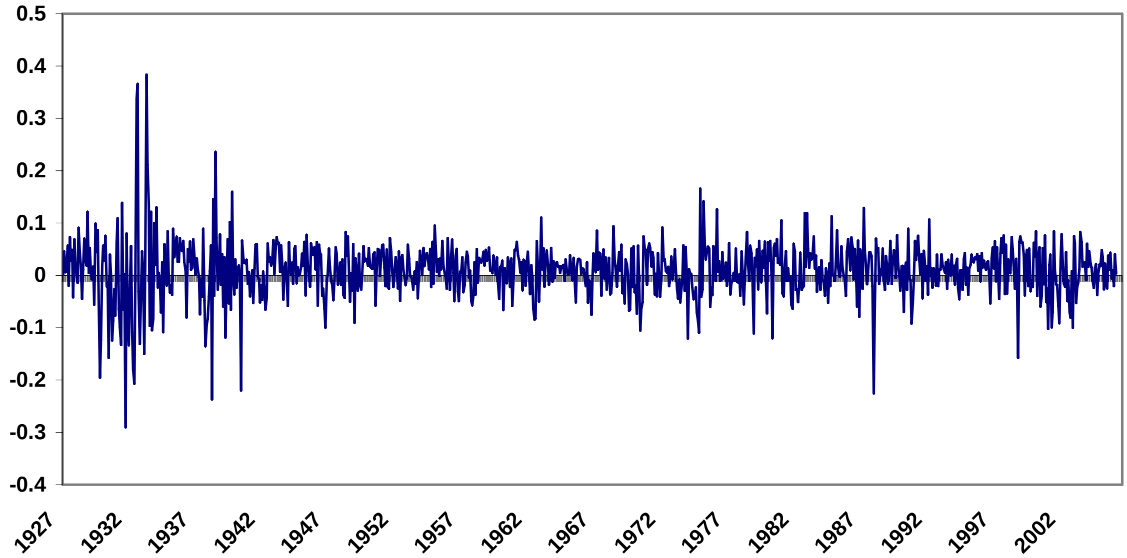


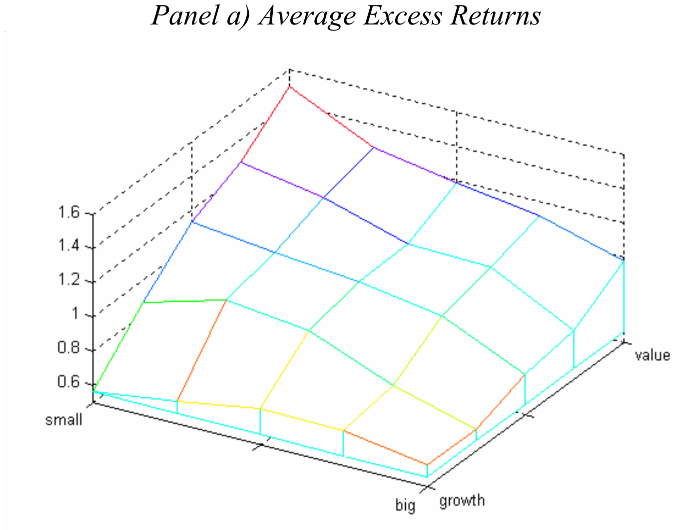
Figure 3
Average Excess Returns and Volatility Loadings

Factors' loadings are estimated using a model which has an *AR(1)-IGARCH(1,1)* specification:

$$\begin{cases} R_{i,t}^{e} = \alpha_i + \beta_i^{m} R_{m,t}^{e} + \beta_i^{IV} IV_t + \eta_{i,t} \\ \eta_{i,t} = \varepsilon_{i,t} - \gamma_i \eta_{i,t-1} \\ \varepsilon_{i,t} = \sqrt{h_{i,t}}\, e_{i,t} \qquad\qquad i = 1,\ldots,25;\ t = 1,\ldots,T \\ h_{i,t} = \omega_i + \theta_{1,i}\varepsilon_{i,t-1}^{2} + \phi_{1,i} h_{i,t-1} \\ e_{i,t} \sim \mathrm{N}(0,1) \end{cases}$$

R_m^e represents the excess market return and *IV* is the innovations in ex-post volatility. Data span the period January 1927-December 2005.

Panel a) Average Excess Returns



Panel b) (–Volatility) Loadings $\hat{\beta}_{\mathrm{i}}^{-\mathrm{IV}}$

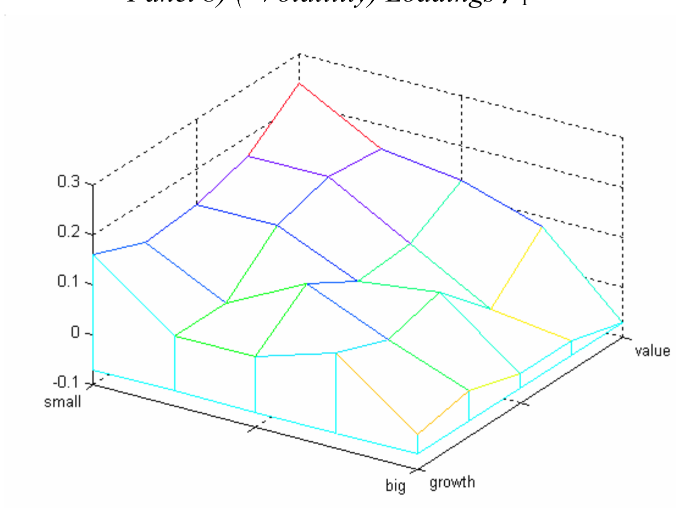


Figure 4
Conditional Volatilities

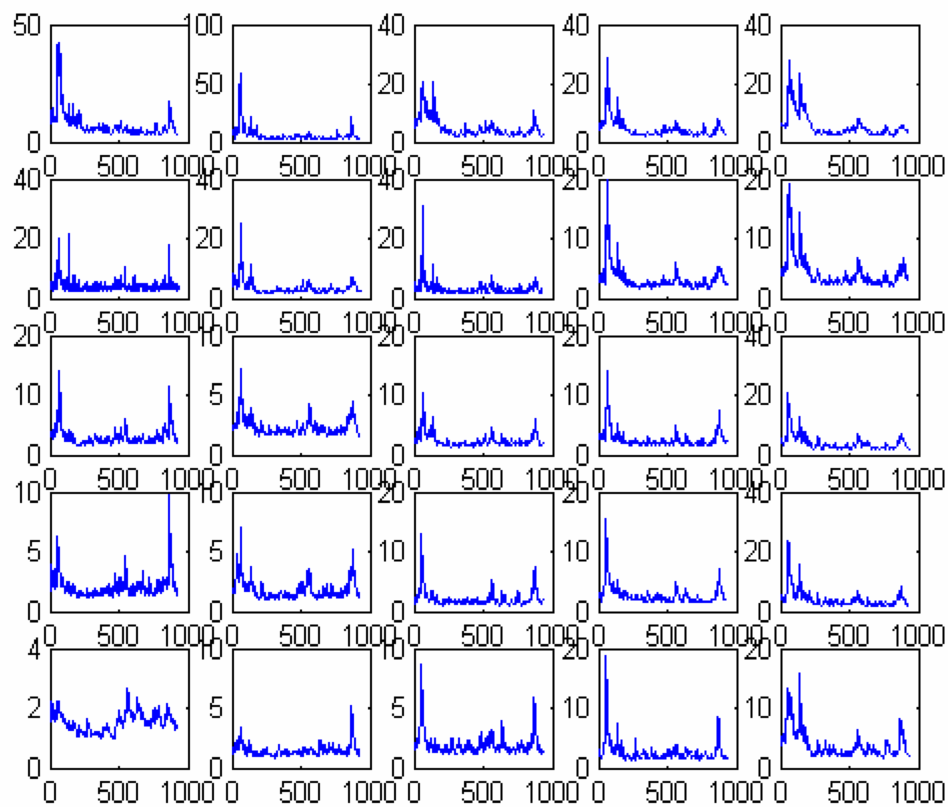
The conditional error standard deviations are estimated for the period January 1927 – December 2005, for the 25 Fama-French portfolios sorted on size and book-to-market equity, using a model which has an *AR(1)-IGARCH(1,1)* specification for the model error:

$$\begin{cases} R_{i,t}^{e} = \alpha_i + \beta_i^{m} R_{m,t}^{e} + \beta_i^{IV} IV_t + \eta_{i,t} \\ \eta_{i,t} = \varepsilon_{i,t} - \gamma_i \eta_{i,t-1} \\ \varepsilon_{i,t} = \sqrt{h_{i,t}}\, e_{i,t} & i = 1,\ldots,25; t = 1,\ldots,T \\ h_{i,t} = \omega_i + \theta_{1,i}\varepsilon_{i,t-1}^{2} + \phi_{1,i} h_{i,t-1} \\ e_{i,t} \sim \mathrm{N}(0,1) \end{cases}$$

The number of time series observations is equal to 933. The rows below index firm size, while the columns index firm value. Conditional volatilities are reported in percentages.

Low BE/ME ***High BE/ME***

Small

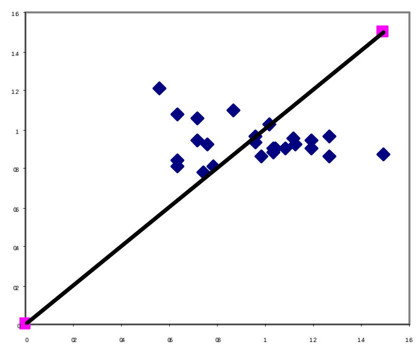


Big

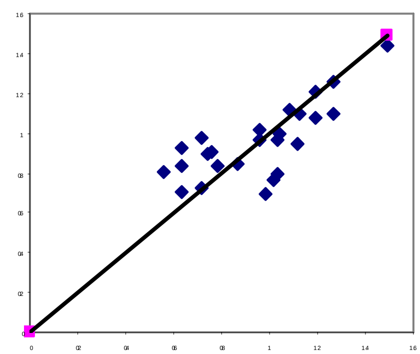
Figure 5
Predicted vs Realized Average Excess Returns

Predicted excess returns are plotted on the vertical axis, while realized excess returns are plotted on the horizontal axis, using 3 different asset pricing models: the CAPM, the 2-factor (consisting of market excess return and innovations in its volatility) and the Fama-French 3 factor model (FF-3). Panel b) assumes that the borrowing and lending rates are different (the intercept α in the cross-sectional regression is different from zero), while Panel c) assumes that the two rates are equal *(α=0)*.

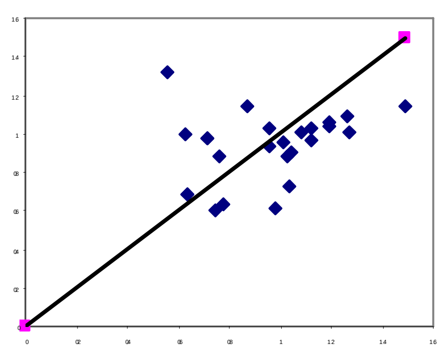
Panel a) CAPM



Panel b) 2-Factor Model with Intercept



Panel c) 2-Factor Model without Intercept



Panel d) FF-3 Factor Model

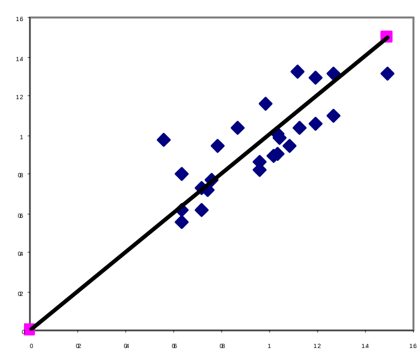


Table I
Estimates for the AR(1)-IGARCH(1,1) Model for the Ex-Post Volatility

Monthly portfolio return data collected for the period January 1927 – December 2005 for the 25 Fama-French portfolios sorted on size and book-to-market equity are obtained from Kenneth French's web-site at Dartmouth. The first measure of stock market volatility is similar to the one used by Anderson et al. (2002) and is given by

$$\nu_{t,t+\Delta} = \sqrt{\sum_{i=1}^{\Delta-1}\left(R_{m,t+i+1} - R_{m,t+i}\right)^2},$$

where $R_{m,t}$ denotes daily return on the stock market portfolio and Δ represents the number of trading days in a given month. $IV_t = \nu_{t,t+\Delta} - \nu_{t-\Delta,t}$ represents innovations in ex-post volatility. The time series of portfolios excess returns *($R^e_{i,t}$)* are regressed on the excess market return *($R^e_{m,t}$)* and IV_t in a model in which the error follows an *AR(1)-IGARCH(1,1)* specification:

$$\begin{cases} R^e_{i,t} = \alpha_i + \beta_i^m R^e_{m,t} + \beta_i^{IV} IV_t + \eta_{i,t} \\ \eta_{i,t} = \varepsilon_{i,t} - \gamma_i \eta_{i,t-1} \\ \varepsilon_{i,t} = \sqrt{h_{i,t}}\, e_{i,t} \qquad\qquad i = 1,...,25; t = 1,...,T \\ h_{i,t} = \omega_i + \theta_{1,i}\varepsilon^2_{i,t-1} + \phi_{1,i} h_{i,t-1} \\ e_{i,t} \sim \mathrm{N}(0,1) \end{cases}$$

The estimation is done using the maximum likelihood method and the initial values for the regression parameters are obtained from the Yule-Walker estimates. The *t*-statistics are reported in parentheses.

Size Quintiles	Book-to-Market Equity (BE/ME) Quintiles									
	Low	2	3	4	High	Low	2	3	4	High
					$\hat{\alpha}_i$					
Small	-0.50	-0.11	0.18	0.37	0.73	-(2.98)	-(0.86)	(1.53)	(3.21)	(4.91)
2	-0.12	0.06	0.26	0.39	0.48	-(0.89)	(0.58)	(3.21)	(4.23)	(4.17)
3	-0.13	0.16	0.22	0.34	0.41	-(1.54)	(2.12)	(3.10)	(4.44)	(3.57)
4	-0.03	0.03	0.21	0.21	0.34	-(0.54)	(0.57)	(3.45)	(2.80)	(3.29)
Big	-0.04	-0.02	0.17	0.11	0.16	-(0.87)	-(0.40)	(2.92)	(1.47)	(1.79)
					$\hat{\beta}_i^m$					
Small	1.37	1.25	1.05	0.95	0.94	(40.45)	(52.87)	(61.81)	(52.79)	(44.91)
2	1.24	1.18	1.02	1.01	1.09	(55.78)	(140.47)	(68.88)	(62.14)	(57.62)
3	1.22	1.09	1.03	1.03	1.06	(84.74)	(79.67)	(85.36)	(77.27)	(53.75)
4	1.08	1.07	1.06	1.03	1.09	(87.27)	(118.81)	(103.78)	(67.80)	(55.79)
Big	0.98	0.94	0.91	0.95	1.02	(122.38)	(110.45)	(85.80)	(65.98)	(48.07)
					$\hat{\beta}_i^{IV}$					
Small	-0.16	-0.12	-0.13	-0.16	-0.24	-(2.31)	-(2.26)	-(3.27)	-(5.18)	-(5.55)
2	-0.04	-0.04	-0.13	-0.16	-0.15	-(1.23)	-(1.70)	-(4.68)	-(5.34)	-(3.83)
3	-0.04	-0.12	-0.06	-0.07	-0.13	-(1.44)	-(4.60)	-(2.74)	-(3.22)	-(3.38)
4	-0.09	-0.05	-0.08	0.02	-0.08	-(4.39)	-(3.39)	-(3.90)	(0.97)	-(2.63)
Big	0.03	0.01	0.04	0.04	0.07	(1.78)	(0.28)	(1.82)	(1.98)	(2.03)

Table I - *Continued*

Size Quintiles	Low	2	3	4	High	Low	2	3	4	High
	Book-to-Market Equity (BE/ME) Quintiles									
					$\hat{\gamma}_i$					
Small	-0.05	-0.04	-0.11	-0.11	-0.18	-(1.22)	-(1.17)	-(3.17)	-(3.27)	-(5.30)
2	-0.12	-0.08	-0.08	-0.09	-0.09	-(2.96)	-(2.13)	-(2.09)	-(2.56)	-(2.42)
3	-0.04	-0.04	-0.06	-0.03	-0.14	-(1.18)	-(1.04)	-(1.78)	-(0.83)	-(3.67)
4	-0.03	-0.05	-0.08	-0.07	-0.10	-(0.76)	-(1.54)	-(2.10)	-(1.89)	-(2.80)
Big	-0.10	-0.01	-0.05	-0.01	-0.03	-(3.15)	-(0.16)	-(1.18)	-(0.22)	-(0.81)
					$\hat{\omega}_i$					
Small	0.56	0.82	0.16	0.20	0.03	(2.61)	(3.92)	(2.94)	(3.42)	(0.74)
2	3.65	0.18	0.60	0.19	0.25	(7.36)	(2.70)	(5.00)	(2.89)	(3.19)
3	0.23	0.16	0.16	0.24	0.15	(3.16)	(2.72)	(3.75)	(3.36)	(2.42)
4	0.22	0.10	0.14	0.17	0.29	(3.16)	(3.21)	(3.15)	(3.53)	(3.70)
Big	0.04	0.11	0.21	0.39	0.26	(2.37)	(3.46)	(4.21)	(7.62)	(4.04)
					$\hat{\theta}_{1,i}$					
Small	0.14	0.24	0.13	0.12	0.08	(12.55)	(11.97)	(9.51)	(9.21)	(15.05)
2	0.35	0.16	0.29	0.12	0.11	(12.00)	(7.77)	(11.54)	(8.17)	(8.69)
3	0.15	0.09	0.11	0.12	0.13	(7.18)	(5.18)	(6.64)	(7.24)	(11.71)
4	0.15	0.14	0.18	0.13	0.16	(7.74)	(6.82)	(6.90)	(10.37)	(10.39)
Big	0.05	0.12	0.14	0.18	0.12	(4.07)	(5.34)	(6.40)	(7.22)	(8.47)
					$\hat{\phi}_{1,i}$					
Small	0.86	0.76	0.87	0.88	0.93	(79.47)	(42.90)	(75.45)	(68.72)	(328.37)
2	0.48	0.85	0.68	0.87	0.88	(11.78)	(45.85)	(25.22)	(51.87)	(62.27)
3	0.84	0.88	0.87	0.84	0.87	(39.90)	(38.20)	(48.83)	(39.45)	(91.55)
4	0.80	0.84	0.80	0.86	0.83	(24.64)	(39.49)	(27.66)	(59.47)	(64.93)
Big	0.93	0.83	0.81	0.77	0.86	(52.97)	(27.91)	(27.42)	(32.71)	(59.06)

R^2	*Book-to-Market Equity (BE/ME) Quintiles*				
Size Quintiles	***Low***	***2***	***3***	***4***	***High***
Small	0.51	0.54	0.63	0.63	0.56
2	0.71	0.76	0.76	0.75	0.69
3	0.82	0.86	0.85	0.80	0.73
4	0.87	0.90	0.86	0.81	0.72
Big	0.92	0.91	0.85	0.77	0.66

Table II
Estimates for Factors' Risk Prices

Factors' risk prices for the period January 1927-December 2005 in the following asset pricing model $E\left(R_{i,t+1}^{e}\right)=\lambda_m\beta_i^m+\lambda_{IV}\beta_i^{IV}$ are estimated using the Fama-MacBeth (1973) procedure. The left-hand side variable is a vector of size- and value-sorted portfolios expected excess returns, and the betas represent factors' loadings estimated in the corresponding time series models. I assume that the borrowing and lending rates are equal (no intercept in the cross-sectional regressions). *IV* represents innovations in ex-post volatility. The *t*-statistics are reported in parentheses. R^2s are redefined.

$\hat{\lambda}_m$	$\hat{\lambda}_{IV}$	*Avg MSE*	R^2	*Adj* R^2
0.74	-1.90	10.35	0.94	0.93
(3.83)	-(3.04)			

Table III
Estimates for Volatility Risk Premia

The risk premia for small stocks versus big stocks are estimated as the product between ex-post volatility risk price and the differences in the factors' loadings between small and big stocks, using a 2-factor model consisting of the return on the market portfolio and innovations in its volatility. Results are reported in percentages, on a monthly basis.

	Book-to-Market Equity (BE/ME) Quintiles				
	Low	***2***	***3***	***4***	***High***
$\hat{\lambda}_{IV}(\hat{\beta}_{Small}^{IV}-\hat{\beta}_{Big}^{IV})$	0.36	0.25	0.32	0.38	0.59

Table IV
Estimates for Factors' Risk Prices – Robustness with Respect to Other Risk Factors

Factors' risk prices for the period January 1927-December 2005 are estimated using the Fama-MacBeth (1973) procedure for the following asset pricing model:

$$E\left(R_{i,t+1}^{e}\right)=\lambda_m \beta_i^m+\sum_{s=1}^{p}\lambda_s \beta_i^s ,$$

where λ_m is the price of market risk and λ_s is the price of risk associated with the generic factor *s*. The left-hand side variable is a vector of size- and value-sorted portfolios expected excess returns, and the betas represent factors' loadings estimated in the corresponding time series models. I assume that the borrowing and lending rates are equal (no intercept in the cross-sectional regressions). *IV* represents innovations in ex-post volatility. The *t*-statistics are reported in parentheses. R^2s are redefined.

$\hat{\lambda}_m$	$\hat{\lambda}_{IV}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	*Avg MSE*	R^2	R^2_{adj}
0.86 (4.05)					11.85	0.92	0.92
0.74 (3.83)	-1.90 -(3.04)				10.35	0.94	0.93
0.68 (3.80)		0.21 (1.55)	0.49 (4.15)		7.12	0.98	0.98
0.68 (3.81)	-2.11 -(4.27)	0.24 (1.88)	0.52 (4.47)		6.98	0.99	0.98
0.69 (3.92)	-2.11 -(3.44)	0.25 (1.93)	0.53 (4.59)	0.65 (0.90)	6.93	0.99	0.98

Table V
Estimates for Factors' Risk Prices – Different Samples

Factors' risk prices are estimated using the Fama-MacBeth (1973) procedure for the following asset pricing model:

$$E\left(R_{i,t+1}^{e}\right)=\lambda_m\beta_i^m+\sum_{s=1}^{p}\lambda_s\beta_i^s \ ,$$

where λ_m is the price of market risk and λ_s is the price of risk associated with the generic factor *s*. Two additional sample periods are used. The left-hand side variable is a vector of size- and value-sorted portfolios expected excess returns, and the betas represent factors' loadings from the corresponding time series models. I assume that the borrowing and lending rates are equal (no intercept in the cross-sectional regressions). *IV* represents innovations in ex-post volatility. The *t*-statistics are reported in parentheses. R^2s are redefined.

Panel a) Post-depression sample (January 1935-December 2005)

$\hat{\lambda}_m$	$\hat{\lambda}_{IV}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	*Avg MSE*	R^2	R^2_{adj}
0.84					7.14	0.91	0.91
(4.71)							
0.70	-2.03				5.91	0.93	0.93
(4.32)	-(3.65)						
0.68		0.17	0.49		3.58	0.98	0.97
(4.37)		(1.42)	(4.62)				
0.69	-2.10	0.20	0.51		3.32	0.99	0.98
(4.38)	-(4.51)	(1.81)	(4.82)				
0.69	-2.05	0.21	0.52	0.49	3.30	0.99	0.98
(4.40)	-(3.59)	(1.83)	(4.85)	(0.85)			

Table V - *Continued*

Panel b) Post-COMPUSTAT sample (January 1963-December 2005)

$\hat{\lambda}_m$	$\hat{\lambda}_{IV}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	*Avg MSE*	R^2	R^2_{adj}
0.66					5.98	0.85	0.85
(3.17)							
0.52	-1.23				4.55	0.87	0.86
(2.72)	-(2.29)						
0.47		0.23	0.51		2.41	0.97	0.97
(2.42)		(1.55)	(3.70)				
0.46	-1.23	0.25	0.50		2.16	0.97	0.97
(2.39)	-(2.27)	(1.74)	(3.66)				
0.48	-1.15	0.26	0.52	1.77	2.13	0.98	0.97
(2.49)	-(2.05)	(1.80)	(3.86)	(2.73)			

Table VI
Estimates for Factors' Risk Prices – Robustness with Respect to Model Specification

Factors' risk prices are estimated using the Fama-MacBeth (1973) procedure for the period January 1927-December 2005 for the following asset pricing model

$$E\left(R_{i,t+1}^{e}\right)=\alpha+\lambda_{m}\beta_{i}^{m}+\sum_{s=1}^{p}\lambda_{s}\beta_{i}^{s} ,$$

where λ_m is the price of market risk and λ_s is the price of risk associated with the generic factor *s*. The left-hand side variable is a vector of size- and value-sorted portfolios expected excess returns, and the betas represent factors' loadings estimated in the corresponding time series models. I assume that the borrowing and lending rates are different, thus the intercept in the model. *IV* represents innovations in ex-post volatility. The *t*-statistics are reported in parentheses. R^2s are redefined.

α	$\hat{\lambda}_m$	$\hat{\lambda}_{IV}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	*Avg MSE*	R^2	R^2_{adj}
1.49	-0.50					10.51	0.05	0.01
(3.57)	-(1.06)							
1.94	-1.07	-2.11				9.28	0.58	0.54
(5.20)	-(2.79)	-(3.44)						
1.63	-0.90		0.20	0.48		6.88	0.80	0.77
(4.33)	-(2.20)		(1.48)	(4.13)				
1.38	-0.65	-1.49	0.23	0.51		6.73	0.87	0.84
(3.71)	-(1.63)	-(2.92)	(1.78)	(4.38)				
1.41	-0.68	-1.61	0.23	0.51	-0.14	6.67	0.87	0.84
(3.56)	-(1.65)	-(2.70)	(1.81)	(4.43)	-(0.18)			

Table VII
Estimates for Factors' Risk Prices – Robustness with Respect to Alternative Volatility Measures

Factors' risk prices are estimated using the Fama-MacBeth (1973) procedure for the period January 1927-December 2005 for the following asset pricing model:

$$E\left(R^{e}_{i,t+1}\right) = \lambda_m \beta_i^m + \sum_{s=1}^{p} \lambda_s \beta_i^s \ ,$$

where λ_m is the price of market risk and λ_s is the price of risk associated with the generic factor s. The left-hand side variable is a vector of size- and value-sorted portfolios expected excess returns, and the betas represent factors' loadings estimated in the corresponding time series models. *IV* represents innovations in ex-ante volatility. Ex-ante volatility is built using the estimates from the following model:

$$\begin{cases} R_{m,t} = \mu + \rho R_{m,t-1} + \eta_t \\ \eta_t = \varepsilon_t \sqrt{h_t} \\ \varepsilon_t / I_{t-1} \sim \text{Student}(0,1,\upsilon) \\ h_t = \alpha + \phi h_{t-1} + \theta^{+}\eta^{2+}_{t-1} + \theta^{-}\eta^{2-}_{t-1} \\ \eta^{2+}_t = \eta^2_t 1_{\{\eta_t > 0\}}, \eta^{2-}_t = \eta^2_t 1_{\{\eta_t < 0\}} \end{cases} \quad , \text{t=1… T.}$$

The *t*-statistics are reported in parentheses.

$\hat{\lambda}_m$	$\hat{\lambda}_{IV}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	*Avg MSE*	R^2	R^2_{adj}
0.66					5.98	0.85	0.85
(3.17)							
0.92	-2.75				12.11	0.94	0.94
(4.09)	-(3.19)						
0.68	-2.83	0.22	0.44		6.51	0.98	0.98
(3.79)	-(2.47)	(1.65)	(3.75)				
0.73	-2.76	0.23	0.48	1.93	6.42	0.99	0.98
(4.12)	-(2.33)	(1.75)	(4.16)	(3.16)			

Table VIII
Posterior Estimates for the *Asymmetric Student-GARCH(1,1)* Model

The model used is $\begin{cases} R_{m,t} = \mu + \rho R_{m,t-1} + \eta_t \\ \eta_t = \varepsilon_t \sqrt{h_t} \\ \varepsilon_t / I_{t-1} \sim \text{Student}(0,1,\upsilon) \\ h_t = \alpha + \phi h_{t-1} + \theta^+ \eta_{t-1}^{2+} + \theta^- \eta_{t-1}^{2-} \\ \eta_t^{2+} = \eta_t^2 1_{\{\eta_t > 0\}}, \eta_t^{2-} = \eta_t^2 1_{\{\eta_t < 0\}} \end{cases}$ *, t=1… T.*

R_t represents the monthly time series of returns on the stock market portfolio for the period 1962 - 2001. The results are computed using a Griddy-Gibbs sampling algorithm in which 1000 draws are kept and 1000 initial draws are the burn-ins sample. A flat prior on finite intervals is used on all parameters except for the prior on υ, which is half-Cauchy. Standard errors are reported in the parentheses.

	Estimate	***Std Err***
μ	0.20	(0.03)
ρ	0.05	(0.03)
α	0.09	(0.04)
θ^+	0.03	(0.02)
ϕ	0.75	(0.07)
υ	7.90	(2.78)
θ^-	0.18	(0.06)

Table IX
Average Excess Returns for the 25 Fama-French Portfolios

Monthly average excess returns for the 25 size- and value-sorted portfolios are reported in percentages. *T*-statistics are documented in parentheses.

	Book-to-Market Equity (BE/ME) Quintiles									
Size Quintiles	**Low**	**2**	**3**	**4**	**High**	**Low**	**2**	**3**	**4**	**High**
Small	0.56	0.87	1.13	1.27	1.50	(1.39)	(2.45)	(3.81)	(4.50)	(4.76)
2	0.63	1.01	1.08	1.19	1.27	(2.37)	(3.87)	(4.47)	(4.79)	(4.45)
3	0.71	0.96	1.03	1.04	1.19	(2.80)	(4.45)	(4.63)	(4.69)	(4.25)
4	0.71	0.76	0.96	1.03	1.12	(3.48)	(3.72)	(4.57)	(4.53)	(3.82)
Big	0.63	0.63	0.74	0.78	0.98	(3.52)	(3.68)	(3.96)	(3.45)	(3.90)

Table X
Descriptive Statistics

R^e_m is the excess market return. *MOM* is the momentum factor of Jegadeesh and Titman (1993). *HML* and *SMB* are mimicking portfolios for book-to-market equity and size (zero-investment portfolios). *HML* is the difference between high book-to-market-stocks portfolios and low book-to-market-stocks portfolios, with similar weighted-average size, while *SMB* is the difference between the returns on small-stocks portfolios and those of big-stocks portfolios, with similar weighted-average book-to-market equity. *IV* represents innovations in ex-post volatility.

Panel a) Factors' Means

	R^e_m	SMB	HML	MOM
Average	0.65	0.25	0.41	0.76

Panel b) Cross-Correlation Matrix

	R^e_m	IV	SMB	HML	MOM
R^e_m	1.00				
IV	-0.30	1.00			
SMB	0.33	-0.19	1.00		
HML	0.20	0.00	0.10	1.00	
MOM	-0.31	-0.05	-0.17	-0.38	1.00